

Summer 2019

## **Leveraging Geotagged Social Media to Monitor Spatial Behavior During Population Movements Triggered by Hurricanes**

Yago Martín González

Follow this and additional works at: <https://scholarcommons.sc.edu/etd>



Part of the [Geography Commons](#)

---

LEVERAGING GEOTAGGED SOCIAL MEDIA TO MONITOR SPATIAL BEHAVIOR  
DURING POPULATION MOVEMENTS TRIGGERED BY HURRICANES

by

Yago Martín González

Bachelor of Arts  
University of Zaragoza, 2014

Master of Arts  
University of Zaragoza, 2015

---

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Geography

College of Arts and Sciences

University of South Carolina

2019

Accepted by:

Susan L. Cutter, Major Professor

Zhenlong Li, Committee Member

Jerry T. Mitchell, Committee Member

Christopher T. Emrich, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

© Copyright by Yago Martín González, 2019  
All Rights Reserved.

## **DEDICATION**

Not even a thousand dedicated dissertations would suffice to thank you for being such a brave and loving mother. What you have done for me is unpayable and it is beyond what words can describe. You always put my sister and me first, and despite the adversities, you managed to pull us all through.

This is as much yours as it is mine. Thank you, my hero.

## **ACKNOWLEDGMENTS**

First, I would like to express my endless gratitude to those closest to me. To my family, friends, and dog for always being there for me. This would had never been possible without your love and support. Special thanks to those friends made during this doctoral journey, my tennis team, and to all the HVRI crew. Particular thanks to Raelene, my severe weather expert, original officemate, and dedicated kickball coach. Tyler, for making me fall in love with what you call “football” and for unsuccessfully teaching me how to throw (my bad!). Sahar, for being the sweetest hazard in the lab. And, Erika, for all your patience, language edits, and driving me across the continent on a trip I will never forget. Y’all have a special place in my heart and I will truly miss you.

Second, all my gratitude and admiration to Susan Cutter for all the support, flexibility, and guidance, for your patience and numerous edits, and for pushing me when needed. You have always done so with a positive attitude and sense of humor. Thank you also to Zhenlong Li for always being available for a chat about my “data struggles” and constantly coming up with the right solution. Thank you as well to the rest of my committee for valuable comments and improvements to my dissertation.

Finally, thank you to Fulbright and all who believed in me to succeed in this wonderful experience. I can’t appreciate enough the huge investment that the program, my country, and the sponsors have made in my personal and professional development.

## **ABSTRACT**

In a world of increased mobility and interconnectedness, the study of spatial behavior becomes more relevant than ever. However, multiple researchers have highlighted that the understanding of these dynamic processes has reached a bottleneck derived from the rigidity of traditional spatial behavior inquiry methods and the unavailability of trustworthy and relevant information. These difficulties are even more prominent during emergencies and disasters as these events often create scenarios where spatial behavior does not follow regular and logical patterns and where conventional mobility datasets are often skewed or not existent. Thus, many scholars working within the spatial behavior sub-discipline are pursuing innovative data collection methods to deepen the understanding of human spatial behavior. Researchers see digital geospatial trace data, also known as passive citizen sensor data, as one of the most promising opportunities to develop and test new hypotheses on spatial behavior. Nevertheless, the application of these new methods has not been fully explored within the hazard/disaster discipline for spatial behavior purposes under stressed situations.

This dissertation investigates the suitability of geotagged social media (Twitter) as an innovative approach for the study of spatial behavior of people in stressed contexts and responds to three main research questions: 1) How well do geotagged social media estimate hurricane evacuation compliance? 2) To what extent is geotagged social media amenable for determining hurricane evacuation behavior? 3) How suitable is geotagged social media

to evaluate post-disaster displacement and tourist flows? The dissertation therefore not only attempts to develop a new method to estimate the number of movements associated with the different stages of an emergency but also tries to answer long-standing questions about the response of different population sub-groups (residential status, gender, age, race/ethnicity) before, during, and after hurricanes. Results confirm the potential of geotagged social media to tackle some of the deficiencies of traditional approaches, particularly offering more timely, dynamic, and affordable information about the evacuation and post-disaster population movements. In addition, results demonstrate that the Twitter-based approach complements survey-based methods as it permits accessing underrepresented groups in traditional approaches such as the young, short-term residents, and racial/ethnic minorities. Although the representativeness of Twitter samples is still debatable and needs further research, this method to investigate emergency-triggered population movements can ultimately improve our understanding of the response and recovery phases of a disaster.

## TABLE OF CONTENTS

DEDICATION .....	iii
ACKNOWLEDGMENTS .....	iv
ABSTRACT.....	v
LIST OF TABLES .....	ix
LIST OF FIGURES .....	x
CHAPTER 1. INTRODUCTION .....	1
1.1. A SHRINKING WORLD .....	2
1.2. THE STUDY OF SPATIAL MOBILITY: BEHAVIORAL GEOGRAPHY .....	5
1.3. KEEPING UP WITH SPACE-TIME COMPRESSION: GEOSPATIAL DIGITAL TRACE DATA .....	10
CHAPTER 2. LEVERAGING TWITTER TO GAUGE EVACUATION COMPLIANCE: SPATIOTEMPORAL ANALYSIS OF HURRICANE MATTHEW .....	18
2.1 INTRODUCTION.....	19
2.2 SOCIAL MEDIA USAGE IN DISASTERS .....	21
2.3. HURRICANE MATTHEW AND STUDY AREA .....	25
2.4. METHODS: DATA COLLECTION AND PREPROCESSING .....	26
2.5. TWITTER REACTION TO HURRICANE MATTHEW .....	29
2.6. RESPONSE TO THE EVACUATION ORDER .....	33
2.7. LIMITATIONS .....	38
2.8. CONCLUSIONS .....	40



CHAPTER 3. BRIDGING TWITTER AND SURVEY DATA FOR THE EVACUATION ASSESSMENT OF HURRICANE MATTHEW AND HURRICANE IRMA.....	54
3.1 INTRODUCTION.....	55
3.2 RESEARCH CONTEXT AND BACKGROUND .....	56
3.3. DISASTER CONTEXT AND STUDY AREAS .....	64
3.4. METHODS.....	65
3.5. RESULTS AND DISCUSSION .....	69
3.6. CONCLUSIONS AND FURTHER RESEARCH .....	77
CHAPTER 4. TRACKING THE DISRUPTION OF HURRICANE MARIA ON POPULATION MOVEMENTS IN PUERTO RICO THROUGH GEOTAGGED TWEETS.....	91
4.1 INTRODUCTION.....	92
4.2 RESEARCH CONTEXT AND BACKGROUND .....	94
4.3. DISASTER CONTEXT AND STUDY AREA .....	103
4.4. METHODS.....	106
4.5. RESULTS AND DISCUSSION .....	108
4.6. CONCLUSIONS AND FURTHER RESEARCH .....	114
4.7. ACKNOWLEDGEMENTS .....	116
CHAPTER 5. CONCLUSIONS .....	124
5.1. TRACKING GEOSPATIAL DIGITAL FOOTPRINTS: AN OPPORTUNITY	125
5.2. A GLANCE AT THE FUTURE .....	132
REFERENCES .....	134
APPENDIX A. PLOS ONE LICENSE .....	168

## LIST OF TABLES

Table 2.1. Twitter datasets .....	45
Table 2.2. Evacuation rates and estimated evacuees by conglomerate.....	51
Table 2.3. Evacuation destinations from South Carolina during Hurricane Matthew .....	52
Table 2.4. Return after evacuation .....	53
Table 3.1. Survey response rates and confidence interval .....	80
Table 3.2. Dataset query conditions.....	81
Table 3.3. Twitter active users in the study areas .....	81
Table 3.4. Race/ethnic categories .....	82
Table 3.5. Dependent and independent variables for evacuation behavior analysis .....	82
Table 3.6. Demographic profile of survey and twitter samples (Hurricane Matthew). ....	83
Table 3.7. Statistical significance of predictors of evacuation behavior for Hurricane Matthew in the coastal counties of South Carolina .....	84
Table 3.8. Evacuation rates and odds ratio for Hurricane Matthew in the coastal counties of South Carolina .....	85
Table 3.9. Statistical significance of predictors of evacuation behavior for Hurricane Irma in Florida .....	87
Table 3.10. Evacuation rates and odds ratio for Hurricane Irma in Florida .....	89
Table 4.1. Travel duration.....	119
Table 4.2. Destination of displaced.....	120
Table 4.3. Chi-Squared tests of independence.....	121
Table 4.4. Displacement rates and odds ratio for different age groups .....	122

## LIST OF FIGURES

Figure 1.1. Behavioral geography in the context of the social sciences. ....	17
Figure 2.1. Hurricane Matthew and the study area. ....	43
Figure 2.2. Overview of the research approach and the Twitter datasets used.....	44
Figure 2.3. Temporal distribution of Matthew-related tweets for the five states .....	46
Figure 2.4. Temporal distribution of evacuation-related tweets. ....	47
Figure 2.5. Spatial distribution of Matthew-related Twitter activity. ....	48
Figure 2.6. Correlation between Twitter activity per county and distance to Hurricane Matthew. ....	49
Figure 2.7. Workflow to obtain the local users.....	50
Figure 2.8. Response to the evacuation order of the 1384 local Twitter users. ....	51
Figure 2.9. Evacuation destinations from South Carolina during Hurricane Matthew. ...	52
Figure 3.1. Area of study and storms tracks. ....	80
Figure 3.2. Comparison of age and gender distribution between survey and Twitter samples and the overall population of the coastal counties of South Carolina.....	84
Figure 3.3. Evacuation behavior by age group and method. ....	87
Figure 3.4. Comparison of age and gender distribution between Twitter sample and the overall population of the studied counties of Florida .....	88
Figure 4.1. Hurricane Maria in Puerto Rico and population forecasts .....	117
Figure 4.2. Workflow to obtain displacement behavior of Puerto Rican residents .....	118
Figure 4.3. Population pyramid of Puerto Rico and Twitter sample. ....	119
Figure 4.4. Timing of departure and return of the displaced .....	121

Figure 4.5. 3-week moving average of non-residents Twitter users active in Puerto Rico .....	122
Figure 4.6. Percentage difference of non-resident Twitter users active in Puerto Rico .	123
Figure A.1. Permission to reprint: Plos ONE license. ....	168

## **CHAPTER 1**

### **INTRODUCTION**

### 1.1. A SHRINKING WORLD

Human mobility refers to the ability of the human body to move or shift, whether socially or across the three-dimensional space. In broad terms, human spatial mobility is the capacity of the human being to physically change their location in the geographic space. Humans are spatially mobile by nature, having mobility in their survival toolset. Just as for many other living organisms, mobility is a mechanism to assure the vital functions: nutrition, interaction, and reproduction. However, human spatial mobility has grown beyond this biological perspective and has become a complex social construction. Throughout history, people's capacities to be mobile have changed, turning local spatial connections into global dependencies, and extending the spatiotemporal limits of humanity into areas never imagined.

Mobility technologies have largely evolved since the early nomad migrations and war expeditions on foot. Early on, humans sought technologies to accelerate movement (e.g. the wheel), but for centuries animal power remained as the only terrestrial transportation option available. In water bodies, the wind propelled vessels that helped colonizing distant worlds. While long-distance movement was possible, it was tremendously challenging, fraught with dangers, and only accessible to a privileged few. After many centuries of contained mobility, the invention of the steam engine and the subsequent Industrialization meant an unprecedented revolution in technologies to move people. During the 19<sup>th</sup> century, the steam boat and steam locomotives boosted regional and global commerce and eased population movements such as transoceanic migrations. The construction of a dense railway network permitted a faster and much more affordable terrestrial transportation. Trains monopolized transportation until the arrival of the automobile in the beginning of

the 20<sup>th</sup> century. With the development of road infrastructure, the automobile finally democratized transportation in developed countries, making private transportation possible for most classes of the social strata and becoming a symbol of the modern industrial development. Mobility thereby changed from the occasional privilege of a few to the daily reality of the many. Almost synchronously with the evolution of the automobile, the development of air transportation significantly improved the speed, range and comfort of long-distance travel. Humanity's search for faster and more efficient transportation is never-ending, and many dedicated efforts into tomorrow's transportation revolution. From self-driving cars to magnetic levitation trains, innovations pursue a clear goal, accelerating the pace of life and social interaction by allowing faster physical mobility and/or permitting multitasking while travelling.

The result of the transportation revolution and today's technology of communication is a process known as time-space compression (Harvey, 1989). This theory argues that postmodernity implies an increased sense of physical and virtual connectivity, fueled by the acceleration in the circulation of goods, people, information, and ideas. Time-space compression means a sort of economic, social, and cultural revolution where motion (mobility) integrates the modern identity and where time becomes a commodity. This context of unprecedented movement brings increased complexity and multiple dualisms: positive mobility (business or tourism) versus negative mobility (organized crime, global terrorism, undocumented migration) (Walters, 2006); career or pleasure travel (voluntary) versus survival travel (forced). Some authors have even looked at postmodern mobility in terms of power:

“People who move and act faster, who come nearest to the momentariness of movement, are now the people who rule. And it is the people who cannot move as quickly, and move conspicuously yet the category of people who cannot at will leave their place at all, who are ruled.” (Bauman, 2000: 119-120)

Travelling for survival returns us to one of the basic functions of mobility for many living organisms, subsistence. Whether chased by a predator or fleeing from a fire, animals are not much different than humans, as retreating is a common protective action from an impending threat. The decision-making process behind the move and the mobility capacity is what makes the difference between instinctive escape and meditated temporal or permanent relocation. Increasingly, humans have more information about the threats (hazards) and considerably more mobility capacity to put themselves out of harm's way. The shrinking world is also a reality for “emergency mobilities” (Adey, 2016). Not long ago, the spatial reaction to incoming threats was much different. For instance, thousands perished hunkering down from a powerful hurricane in Galveston in 1900, where warning came too late for the evacuation potential at the time. Today, evacuees are often counted in the millions every time a major hurricane threatens United States coasts, many relocating hundreds of miles away from their homes within hours. Space-time compression therefore permits an exceptional speed of response to life and death situations, which has contributed to the decrease in the number of deaths from foreseeable natural hazards such as hurricanes (Willoughby, 2012). Hazards not only trigger mobility, they can partially or completely suppress it at different scales. For instance, tourist movements significantly fell after the



9/11 terrorist attacks throughout the world (Bonham et al., 2006) and Eyjafjallajökull volcano paralyzed the air traffic in Europe for several days (Lund and Benediktsson, 2011).

A shrunken world where mobility becomes predominant brings additional threats as well. Not only humans become global. For instance, vector-borne diseases can now be spread throughout the world at alarming speed rates as the vector (human or animal) increases its mobility potential. Invasive species are rapidly colonizing many environments and displacing local flora and fauna populations. Ideas can travel farther and faster as well, and terrorist organizations have found in physical and virtual mobility the means to extend and increase their influence and actions.

In short, the increased connectivity promotes a world where everything comes together, where good and bad, positive and negative, problems and solutions, coexist. In such an intertwined, dynamic, and ever-changing context, the study of the many spheres of mobility (physical, virtual, social, etc.) becomes more relevant than ever, which Sheller and Urry named the “mobility turn”, expanding the idea of the “new mobilities paradigm” (Sheller and Urry, 2006).

## 1.2. THE STUDY OF SPATIAL MOBILITY: BEHAVIORAL GEOGRAPHY

The essence of spatial mobility has long been a central geographical subject as it involves the spatial interaction of the human beings with the environment (Cresswell, 2011). However, the study of spatial mobility has also been a theme of interest for many other social sciences (Figure 1.1), with every discipline defining mobility in distinctive ways and placing the emphasis on different aspects. For instance, sociologists define spatial mobility as the subjective or objective inclination to be mobile to reach places where social

activities take place (Colleoni, 2016). Thus, sociology is particularly interested in the potential for spatial mobility (motility) and the social factors that relate to this (Kaufmann, 2017). Economists interested in mobility focus on spatial mobility towards places where activities that generate direct or indirect benefit occur (Faggian and McCann, 2008). Political science has been especially involved with the legal, political, and ethical aspect of transnational mobility (Squire, 2010). History and anthropology concentrate its research in the temporal and cultural development of mobility (Dalakoglou and Harvey, 2012) while psychology investigates perceptions, attitudes and preferences related to spatial mobility (Van Acker et al., 2010).

The study of spatial decision-making especially draws from the interaction among geography, sociology, and psychology and is the foundation of the sub-discipline called behavioral geography (Figure 1.1). The interest for spatial behavior especially flourished in the 1970s as a reaction to the “quantitative turn” that dominated the spatial sciences during the 1950s and 1960s, and several researchers made numerous theoretical and methodological advances (Anderson, 1971; Pred, 1981). In these early works, behavioral geographers already recognized that people are active decision-makers who reason based on certain beliefs about their spatial behavioral choices, criticizing at the same time abstract and deterministic models of spatial behavior such as those based on gravity models to predict human migration flows (Argent, 2017). Behavioralism therefore defends that there are person-to-person differences in the beliefs, passions, feelings, and/or emotions that determine cognition and produce mental images based on factors such as age, sex, ethnicity, socioeconomic status or physical capacity (Argent, 2017). Consequently, the resulting spatial decision-making process is bound to be different for individuals, groups and

societies under different given settings. Much attention then turned towards spatial problems of forgotten populations or socially disadvantaged groups. Populations such as the elderly (Golant, 1972), the children (Acredolo and Evans, 1980), the blind (Thinus-Blanc and Gaunet, 1997), the carless (Koutsopoulos and Schmidt, 1976) or even the drug addicts (Pettway, 1995) are just examples of this focus on different individuals and groups. Distinctions on spatial behavior have also been studied in terms of gender, where Mei-Po Kwan has made important contributions (Kwan, 2000), and race and ethnicity (Ellis et al., 2004). The concept of constraint has also been largely referenced as multiple dimensions of economic, cultural, social, political, legal, moral environments play a relevant role in the interaction of people and space. Of special significance and salience is the aspect of time-space constraint. Hägerstrand (1970) advanced the bounded region of time and space and advocated for the simultaneous study of time and space spheres. The time-space constraint concept continues to be employed by multiple researchers in different fields (Sasaki and Nishii, 2010; Ren et al., 2014).

The physical environment can also produce constraints to the regular spatial decision-making process. For instance, hazards and disasters often alter the coupled human-environment relationship and create scenarios where spatial behavior under normal conditions does not normally apply. Gilbert White pioneered the study of people's attitudes and strategies in hazardous events (White, 1945), which was the first seed of the later fully developed hazard/disaster research field. White and his disciples continued and consolidated the idea of bounded rationality applied to hazards, defending that individuals make decisions with partial information and that these decisions are influenced by a range of sociocultural, emotional, and economic factors (Kates, 1962). Paul Slovic, coming from

an environmental psychology perspective, made significant contributions to the conceptualization of risk perception and how this affects spatial decision making (Slovic, 1987). Drawing from these early conceptual contributions, many scholars commenced to pay attention to the actual spatial response of people to specific hazards.

Temporarily moving out of harm's way is often viewed as one of the most effective protective strategies to cope with the impact of hazards. Researchers have identified risk and uncertainty and motility (capacity to move) as the major drivers of the decision of evacuation from an impending threat or the decision to stay in place. Uncertainty is characterized by incomplete knowledge, whereas risk is associated with the evaluation of uncertain events, and both create a stage where the process of decision making might not be logical or rational (Golledge and Stimson, 1997). Motility is often a product of factors such as vehicle ownership, age, or socioeconomic status (Lindell et al., 2011). Multiple predictors with theoretical influence over risk perception and the ability to evacuate have been extensively studied (Baker, 1991). These include, but are not limited to, risk area (Zhang et al., 2004; Arlikatti et al., 2006); housing characteristics (Whitehead et al., 2000; Lindell et al., 2011); warnings and storm information (Sorensen and Mileti, 1988); demographics (Gladwin and Peacock, 1997; Dow and Cutter, 1998); and hurricane experience (Dow and Cutter, 1998; Brommer and Senkbeil, 2010). Overarching studies such as Baker (1991) and Huang et al. (2016a) systematically collected and analyzed the findings of empirical studies on this matter and are great sources to account for the state of the art.

Medium-term relocation (displacement) and permanent migration are common spatial responses after environmental changes such as climate change or disasters (McLeman and

Gemenne, 2018). Displacement often refers to the involuntary and unforeseen movement of people from their area of residence due to disaster-related impacts while migration is often seen as a more permanent change of place of residence after careful consideration of their economic and social difficulties in an impacted area (Adger et al., 2018). The link between mobility and environmental change has been studied from multiple perspectives. Some have focused on the effects of migrations on the environment (Bilsborrow, 2002). Others put the focus on how —or if— the environment produces migrations (Hunter et al., 2015). New Orleans diaspora following Hurricane Katrina in 2005 is one of the most studied events and concentrates much of the literature on the topic. Several studies investigated the spatial distribution and destination of the evacuees and the displaced and the decision-making process regarding the duration of the relocation (Fussell et al., 2010; Groen and Polivka, 2010). Recently, some have stressed how vulnerable populations can be paradoxically both trapped in a disaster area or be more prone to be displaced (Black et al., 2013), while Esnard et al. (2011) quantified the risk of population displacement in the coastal counties of the southern and southeast United States. Also, other scholars discuss about whether this displacement or migration can be considered an adaptation failure or an adaptation success (Adger et al., 2018).

Traditionally, short-term pre-event relocation (evacuation), medium-term relocation (displacement), and permanent change of residence (migration) have relied either on aggregated and temporally sparse official statistics —such as censuses— or on selective, small-scale observations and surveys. However, the general study of human movements often encounters difficulties in form of the lack of accessible and reliable data, which becomes even more prominent during emergency and post-disaster situations (Laczko,

2015; Rango and Vespe, 2017). For instance, Finch et al. (2010) warned about the lack of reliable estimates of the number of residents fleeing a disaster-stricken area, whether they return, and their destination choices. Focused on hurricane evacuation research, Baker (2009) called for innovation in data-collection methods:

“More creative methods or combinations of data-collection methods need to be employed... If combined with more traditional methods, they could broaden both the scope and depth of understandings about evacuation behavior, while being benchmarked against measures whose reliability is already known.” (Baker, 2009: 209)

The obstacles to track evacuated, displaced, and migrant groups in a shrunken world are even more noticeable. In a few hours, people escaping from a disaster can be thousands of miles away from the area, in some cases crossing national borders. The rate of emergency-triggered mobility is therefore also accelerated and traditional methods are often too rigid to keep pace with such dynamic processes. Some studies have already investigated alternative data sources such as USPS mail recipient and vacancy data (Finch et al., 2010) or Internal Revenue Services (Fussell et al., 2014a). However, these alternatives seem not to tackle the problem completely and many have stressed the necessity for the search and application of more dynamic methods and datasets for the study of spatial behavior.

### 1.3. KEEPING UP WITH SPACE-TIME COMPRESSION: GEOSPATIAL DIGITAL TRACE DATA

Space-time compression is not just a consequence of a physical transportation revolution but also a result of a set of technological advancements in electronics and computation. The digital revolution that began in the 1960s meant a shift from analogue electronic technology to digital electronics, enabling a deep transformation of production

processes, business practices, and the entertainment industry. Later, the propagation of the Internet and the mass production of computers and cell phones –among other digital devices– resulted in the birth of the Information Age.

In this new digital era, the rate at which information is produced and consumed grows exponentially. A very illustrative report from the Independent Expert Advisory Group on a Data Revolution for Sustainable Development in 2014 estimated that more than 90% of the data at the time had been generated in just the last two years (United Nations, 2014). This immense amount of data produced by the use of digital devices and web-based platforms is commonly referred to as “Big Data”, which many characterized as the “3+1 Vs”: Volume, Velocity, Variety, and more recently, Value (Laczko, 2015). Volume is intrinsic to Big Data and refers to the massive scale of data generated through the use of any kind of digital device. Velocity relates to how rapidly these data are created, processed, distributed, and/or consumed. Variety involves the wide array of formats and types of Big Data. For instance, Geospatial Big Data refers to considerably large data sets of spatial data (data having an implicit or explicit association with a location relative to the Earth) and is part of this broad variety. Within Geospatial Big Data there is a great diversity of data types and formats as well. Similar to how our physical activities leave visible and invisible traces behind (fingerprints, bootprints etc.) that can reveal our physical movements, our daily digital activity leaves thousands of records of our presence in the virtual world. Although these data are virtual, a great part of it has associated physical coordinates that determine where in the globe the piece of data was created, processed, or consumed. This is what we call geospatial digital trace data, also known as passive citizen sensor data.

The last “V” represents value. This value is typically understood as financial value, as companies increasingly begin to collect, aggregate, and analyze Big Data to maximize their profits (IBM, 2011). However, Big Data not only hold a financial value. Researchers have also seen the considerable potential of Big Data for scientific purposes. Particularly, social and behavioral scientists have highlighted the opportunities that emerge through the exploitation of Big Data and have initiated numerous research projects in collaboration with experts from engineering and computer sciences (Isaacson and Shoval, 2006; Snijders et al., 2012):

“...it is clear that research into human spatial behavior is conceivably on the brink of a new phase, one in which it will be much easier to gather and collate data on spatial behavior, and where the quality and accuracy of the data obtained will be greater than anything previously recorded” Isaacson and Shoval (2006, p. 181)

When Isaacson and Shoval (2006) suggested that the study of spatial behavior had entered a new phase, they were referring to a phase where the study of population movements would become much more precise and dynamic through the use of geospatial digital trace data. In the last decade, many fields have begun to explore several sources of passively acquired citizen-sensor data such as mobile telephone data, smart card data, WI-FI and Bluetooth data, and social media. Disciplines such as sociology (Amini et al., 2014), public health (Wesolowski et al., 2012), transportation (Çolak et al., 2016), geography (Phithakkitnukoon et al., 2012), and urban studies (Bajardi et al., 2015) have found mobile telephone data an amenable source of spatial information for research purposes. Smart card data from public transit has obvious utility to mobility researchers, especially in urban



environments. Thus, Gong et al. (2017) exploited metro smartcard records to explore spatiotemporal characteristics of intra-urban trips, whereas Tao et al. (2014) conducted a similar study focusing on bus trips. Tracking capital transactions from tourists permitted Sobolevsky et al. (2014) to infer the spatial behavior of people during their holidays in Spain. More recent studies have shown that Wi-Fi and Bluetooth networks can make indoor tracking achievable and relatively accurate (Górak et al., 2016). As well, geotagged social media constitutes a very prolific source of spatial information for a number of disciplines such as public health (Widener and Li, 2014), criminology (Gerber, 2014) and geography (Zhai et al., 2015)

Leveraging Geospatial Big Data in disaster risk reduction has been identified as one of the most promising opportunities to reduce the impact of hazards (Data-Pop Alliance, 2015). However, despite the broad range of applications in other fields, the hazards/disaster community has not yet fully harnessed the potential of geospatial digital trace data in emergency-related spatial behavior. Only epidemiology -biohazards- has systematically collected and analyzed passive citizen-sensor data (mobile phone records) to infer human mobility patterns associated with different infectious diseases (Wesolowski et al., 2014; Vogel et al., 2015). In a post-disaster environment following the Haiti earthquake in 2011, Bengtsson and his team developed a methodology to detect population movements leading to cholera outbreaks by mining call data records (Bengtsson et al., 2011). Similarly, Dobra et al. (2015) created a successful system for detecting emergency events using mobile phone data and revealed inconclusive spatial behavior patterns (very dependent on the particular emergency event). Reflecting upon the challenges and opportunities of mobile phone data, Wesolowski et al. (2016) concluded that data accessibility remains an issue,

and if access is ever granted, it is far from near real time, which is the greatest potential of these data for disaster management purposes. Other sources such as Wi-Fi and Bluetooth networks also encounter difficulties with data accessibility. In one of the few attempts to explore Wi-Fi signals for disaster risk reduction, Moon et al. (2016) achieved an estimation of the location of buried people in a hypothetical building collapse. Access to geotagged social media data is often much easier, especially in publicly available platforms such as Twitter. A few projects have mined Twitter to estimate and characterize human mobility during emergencies. For instance, Wang and Taylor (2014, 2016) disclosed spatiotemporal patterns of human mobility in different disasters by leveraging geotagged tweets, concluding that mobility in urban environments is altered during disaster events and this disturbance and its duration depend upon the characteristics and location of the event. Chae et al. (2014) investigated the 2013 Moore (Oklahoma) tornado and 2012 Hurricane Sandy events and found different spatiotemporal patterns in the location of Twitter users.

As access to other sources of passive citizen-sensor data remains restricted, the exploitation of social media during disasters is increasing (Horita et al., 2013), with some researchers recognizing the potential of geotagged social media in disaster management and particularly in the study of population movements triggered by an emergency (Chen et al., 2016). Despite this increasing interest, there is little understanding as yet on how well such media reflect the actual spatial behavior during an emergency (Steiger et al., 2015).

This dissertation responds to calls for innovative approaches and data collection methods that permit acquiring more timely and dynamic information for the study of human mobility and particularly for human spatial behavior during emergencies. The lack of in-depth studies on how social media represents the actual spatial behavior in disaster

situations serves as well as motivation for this research effort. Thus, resting on the great variety of today's Big Data (third "V"), this dissertation examines the value (fourth "V") of geospatial digital trace data for improving the understanding of emergency-triggered human spatial mobility, which can ultimately contribute to minimize people's exposure to upcoming threats and shed light into the population recovery in the aftermath of a disaster. More specifically, the dissertation investigates the suitability of geotagged social media (Twitter) to help overcome long-known limitations of traditional spatial behavior research approaches and determine if and how this innovative approach can complement these conventional methods for advancing the discipline. The dissertation is structured through the following three research questions:

Chapter 2: How well do geotagged social media estimate hurricane evacuation compliance?

The second chapter is an exploratory analysis that serves as a test for the suitability of geotagged tweets for gauging the evacuation compliance of Twitter users. Focusing on the evacuation of the coastal counties of South Carolina (Jasper, Beaufort, Colleton, Charleston, Dorchester, Berkeley, Georgetown, and Horry) triggered by 2016 Hurricane Matthew, the chapter validates Twitter estimations through comparison with traffic counts offered by the South Carolina Department of Transportation (SCDOT, 2017) and previous behavioral studies of hurricane evacuation in South Carolina conducted through survey questionnaires (Dow and Cutter, 2000; Dow and Cutter, 2002; Cutter et al., 2011).

Chapter 3: To what extent is geotagged social media amenable for determining hurricane evacuation behavior?

The third chapter presents a more in-depth analysis of the Twitter-based evacuation assessment approach and explores the complementation of the proposed approach with survey questionnaires (traditional method). Focusing on 2016 Hurricane Matthew in South Carolina and 2017 Hurricane Irma in Florida, it investigates the depth and value of Twitter data by looking into under-studied aspects of evacuation behavior such as the difference in the evacuation response between permanent and transient residents and the behavior of the young and racial/ethnic minorities.

Chapter 4: How suitable is geotagged social media for evaluating post-disaster displacement and tourist flows?

Chapter 4 examines the suitability of Twitter data for the assessment of the disruption of population movements triggered by 2017 Hurricane Maria in Puerto Rico. The Twitter-based approach estimates the total displacement, destinations, timing, and return of displaced Puerto Ricans Twitter users and includes as well an analysis of the behavior of different subpopulations (gender, age, and region of residence). The behavior of non-resident (visitors) is also monitored during the year after the disaster. Displacement results are contextualized through comparisons with recent studies about these processes in Puerto Rico such as Teralytics (2018), Sutter and Hernandez (2018), Hinojosa et al. (2018), Hinojosa and Melendez (2018), and United States Census Bureau (2018a) while the post-disaster behavior of visiting Twitter users is confronted to pre-disaster baseline levels.

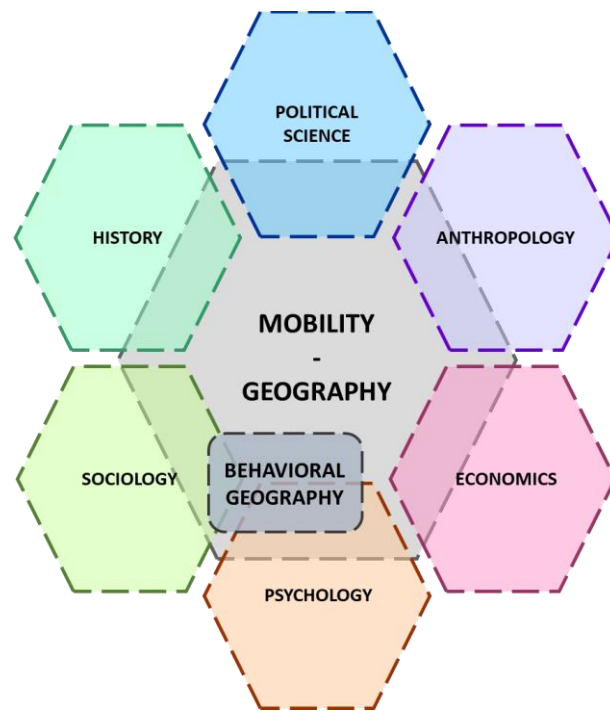


Figure 1.1. Behavioral geography in the context of the social sciences.

## **CHAPTER 2**

### **LEVERAGING TWITTER TO GAUGE EVACUATION COMPLIANCE: SPATIOTEMPORAL ANALYSIS OF HURRICANE MATTHEW<sup>1</sup>**

---

<sup>1</sup> Martín, Y., Li, Z., & Cutter, S. L. (2017). Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew. *PLoS ONE*, 12(7), e0181701. Reprinted here with permission of publisher.

## 2.1 INTRODUCTION

Tropical cyclones represent one of the costliest and devastating threats for populations in both the developed and developing world, with long-lasting consequences that extend over several years (Smith and Katz, 2013; Hsiang and Jina, 2014). The overall profile for natural hazard losses in the U.S. since 1960 finds that tropical cyclones represent roughly 26% of the total losses (Gall et al., 2011), but this percentage has been increasing since 2000 and as of 2015 represents 41% of the total loss from natural hazards (HVRI, 2017). While the projections on frequency and intensity of tropical cyclones remain inconclusive (Knutson et al., 2010), there is increasing risk exposure as population and assets continue to shift to coastal areas (Crossett et al., 2013).

While among the costliest hazards, the death toll from tropical cyclones in the U.S. is relatively low, less than 5% of the hazard fatalities. This is largely due to the protective actions adopted by coastal populations: sheltering in place and evacuation. During tropical cyclones, evacuations are primarily ordered for those who live on the coast or in adjacent low-lying areas due to the anticipated storm surge (Cutter and Smith, 2009). However, the compliance rates with evacuation orders are often significantly less than 100%, with many coastal residents preferring to ride out the storm rather than leave. The decisions to stay or go depend on several factors such as the perception of risk, prior experience, and the severity of the storm itself (Baker, 2000; Dow and Cutter, 2000; Lindell et al., 2005; Bowser and Cutter, 2015). Traditionally, evacuation rates were determined by post-evacuation questionnaire surveys months following the hurricane (Baker, 1991; Dash and Morrow, 2000; Wu et al., 2012; Bukvic et al., 2015). Traffic counts have also been used, but these routinely underestimate the number of people evacuating (Dow and Cutter, 2002)

because the average number of occupants per car is rarely known or estimated. One of the challenges then of emergency management is to assess evacuation order compliance and movements in a timely fashion not only to get people out of harm's way faster, but to facilitate the re-entry into the evacuation zone as quickly as possible once the storm has passed.

Extreme events have the power to attract the public's attention and prompt protective actions. Social media has changed the way we receive information and how we communicate (Kimanuka, 2015). Social media platforms such as Twitter and Facebook offer real-time information and round-the-clock situational awareness and provide a two-way communication channel increasingly used by authorities, businesses, and researchers to quickly grasp public opinion and activity. Today's social media is available on mobile technologies, frequently with built-in Global Positioning System (GPS) receivers. The availability of location-based social networks with geo-tagged social media data has noticeably improved emergency management research and practice (Roick and Heuser, 2013). In fact, emergency and disaster management accounts for 27% of the Twitter data topical applications in reviewed papers (Steiger et al., 2015). Among the many current social media platforms, Twitter is the most often used data source in studies about emergency situations (Sakaki et al., 2011; Mandel et al., 2012; de Albuquerque et al., 2015). Twitter enables the user to create, share, read and reproduce information in a simple and fast manner by limiting the content to 140 characters. These features attract 313 million users per month (Twitter, 2016), producing more than 7,000 tweets per second (Internet Live Stats, 2017).



This paper examines the exchange of hazard information triggered by the most powerful Atlantic Basin hurricane of the 2016 season, Hurricane Matthew. Twitter data are used to gauge the social media reaction to the disaster and the protective actions taken by residents from both a spatial and temporal perspective. More significantly, the paper proposes an approach to leverage social media content to measure the evacuation response in terms of evacuation timing, re-entry, and destinations for safety. Two themes provide the focus for this paper. First, the paper explores the spatiotemporal variability in social media (Twitter) information exchange before, during, and after Hurricane Matthew and whether such differences are due to proximity to the areas of potential impact. The second theme takes a more granular look at South Carolina tweets and examines specific responses to mandatory evacuation orders in South Carolina regarding the number of evacuees, the timing of departure and re-entry, and geographic destination.

## 2.2 SOCIAL MEDIA USAGE IN DISASTERS

Most of the social media usage in a disaster context is focused in the preparedness and response phases of the emergency management cycle. This is to be expected given the real-time need for situational awareness by the public and emergency managers. The concept crisis informatics has been used in the literature to refer to both data and information about emergency response from both official responders and the public. For example, Palen et al. (2009) and Ukkusuri et al. (2014) have helped to further understanding today's crisis informatics, including the use of digital networks (social media) as a mechanism to gauge civil response. In the context of natural hazards, temporal patterns have been identified based on the emergency management cycle (preparedness, response, recovery, mitigation) as well as the affected region or place, the type of event, and characteristics of the event.

For instance, the highest volume of social media content in the Haiyan typhoon was produced days after the landfall (David et al., 2016), illustrating the response/recovery phase. This same pattern was found on another Philippine typhoon, Yolanda in 2013, where the social media activity peak was reached during the response phase, several hours after of the landfall (Meier, 2015). In an earthquake example, Crooks et al. (2013) and Sakaki et al. (2010) leveraged social media response to quickly detect the time and location of the earthquake, essentially using humans as sensors. In a broader study, Huang and Xiao (2015) mined Twitter content to assess the reaction during different disaster phases of Hurricane Sandy, concluding that the activity reached its peak during the initial impact and subsequent hours. This finding was also confirmed by Murthy and Gross (2017) and Kryvasheyev et al. (2015, 2016), who found peak Twitter content on the day of the landfall. Hurricane Irene was also analyzed, finding similar temporal results with the highest number of tweets correlating with the peak of the event (Mandel et al., 2012).

Research using geotagged social media data and identifying spatial patterns of social media users is widespread (Widener and Li, 2014; Gerber, 2014; Zhai et al., 2015). Within disaster management, examples of spatial social media content are found in earthquake, wildfire, tropical cyclone, or flood events (Sakaki et al., 2011; Kent and Capello, 2013; Shelton et al., 2014; Li et al., 2018b). Studies have shown people physically close to a disaster event tend to engage more in disaster-related content in social media. Further, several authors have discussed that when power or phone lines are down or collapsed, Twitter remained active (Li and Rao, 2010; Nguyen et al., 2013). Graham et al. (2013) found through visual interpretation that the highest disaster-related Twitter activity was, as expected, near flooded areas in the United Kingdom during the major floods of 2012.

Introducing the distance to the actual floods, Herfort et al. (2014) and de Albuquerque et al. (2015) proved the same connection between Twitter content and the proximity to a disaster. In the case of tropical cyclones, Kryvasheyev et al. (2016) found a sharp decline in Sandy-related Twitter activity as the distance between the most populated cities in the country and the path of the hurricane increased. In a more focused study on the spatial distribution of Tweets from Hurricane Sandy, Shelton et al. (2014) concluded that the physical distance from the city had no significant relationship with the social media activity about Sandy on Twitter.

Researchers have also been interested in looking at human mobility and social interaction patterns using geotagged social media data. For example, Hasan et al. (2013b) utilize geotagged tweets to identify the spatiotemporal patterns of aggregated and individual mobility in a city. Sadri et al. (2017) develop a modeling approach to explore social interaction networks based on social media interactions. Particularly relevant for our study in the disaster context are the contributions made by Wang and Taylor (2014, 2016), who disclosed spatiotemporal patterns of human mobility in different disasters by leveraging geotagged tweets. They concluded that mobility in urban environments is altered during disaster events and this disturbance and its duration depend upon the characteristics and location of the event. Chae et al. (2014) offered a tool to visually show spatial patterns of Twitter user distribution before, during, and after Hurricane Sandy and the 2013 Moore (Oklahoma) tornado. Building on this scholarship, Chen et al. (2016) theorized about the utility of Volunteering Geographic Information (VGI) such as geotagged tweets in disaster management and in a hypothetical mass evacuation. Our work applies these findings to a case study of an actual hurricane evacuation in order to evaluate

the potential for social media to assist in the quantification of evacuation participation and compliance by residents.

Exploiting social media as a new data stream for evacuation participation addresses some of the issues with the existing methods for behavioral data in hurricane evacuation response. For example, the traditional way of assessing evacuation rates involved surveying the affected population dating back to the early 1960s. While there is a rich historic record of these studies (Baker, 1991), most were funded by federal agencies as part of post-event assessments of evacuations plans. Questionnaire surveys are done by mail and phone but the response rates have been declining. The traditional sampling frames based on address and landline phone are more problematic now because of transition away from landlines to mobile phones and an increasing percentage of unlisted landline numbers. Web-based surveys are cheaper, can be deployed faster, but require internet connectivity and more significantly, user familiarity. Additionally, web surveys can exclude important segments of society (elderly, less educated, the poor) as well as residents in small towns and rural areas (Smyth et al., 2010). Thus, the ability to generalize to a broader population (socially and geographically) is becoming more limiting. Response rates for phone surveys are generally higher (27%) than for mail questionnaire (7%), or web-based surveys (2%) based on a direct comparison in Australia (Sinclair et al., 2012). Direct comparisons of web and mailed questionnaire response rates are quite variable, but on average, mail survey response rates are higher by an average of 20% according to a meta-analysis of US surveys (Shih and Fan, 2009).

Increasingly, states and researchers are looking to integrate traffic data into evacuation planning (Murray-Tuite and Wolshon, 2013). However, such traffic counts are often

devoid of any behavioral assumptions about who is evacuating (in terms of demographic characteristics) and how many people are in each vehicle (Archibald and McNeil, 2012). Thus the count of cars leaving an area may grossly underestimate the number of households or individuals leaving unless calibrated by the number of people per car (or the number of cars per household that would be used in an evacuation). For example, Dow and Cutter (2002) found that one-quarter of the evacuees leaving Charleston in response to Hurricane Floyd (1995) took two or more cars, regardless of household size. By analyzing Twitter users we obtained a sufficiently large dataset to establish spatiotemporal patterns of people's evacuation behaviors without relying on questionnaire surveys or traffic counts.

### 2.3. HURRICANE MATTHEW AND STUDY AREA

Hurricane Matthew was the strongest storm recorded in the Atlantic during the 2016 season, the first Category 5 hurricane since 2007 in this basin (Stewart, 2017). Matthew formed from a tropical wave pushed off the African coast became a named storm September 28th, 2016. Hurricane Matthew rapidly intensified over the warm waters of the Caribbean to a Category 5 storm, and made its first landfall in Haiti as a Category 4 hurricane where its torrential rains caused 546 deaths (Stewart, 2017). On its way north, Matthew made a second landfall in Cuba and a third landfall in the Bahamas causing significant economic damage in both countries. By October 7th, Matthew's track was parallel to the coastline of Florida remaining just offshore as a Category 3 hurricane. A day later, the system started to weaken rapidly as it neared the Georgia coast and made its final landfall (October 8th, 1500 UTC) near McClellanville, South Carolina as Category 1. On October 9th, while still close to the North Carolina coastline, Matthew was downgraded to

an extratropical low, but its windfield and torrential rainfall continued to affect coastal North Carolina and Virginia.

Given the track and early intensity of the storm, coastal residents from Florida to North Carolina were advised to evacuate in advance of the storm. Voluntary and mandatory orders were initially given on October 5th and 6th for the 1.5 million residents living in evacuation zones on the Atlantic coast of Florida. Georgia residents east of Interstate 95 were mandated to evacuate on October 6th. The South Carolina mandatory evacuation order was given by the Governor on October 4th and coastal counties began their evacuation the next day. South Carolina implemented lane reversals on the Interstate 26 to speed up the evacuation.

Figure 2.1 presents the two different study areas established for the research questions. For the spatiotemporal analysis, the entire area affected by Matthew was considered, including the states of Florida (FL), Georgia (GA), South Carolina (SC), North Carolina (NC), and Virginia (VA). In the evacuation behavior analysis, the focus was on the coastal counties of South Carolina (Jasper, Beaufort, Colleton, Charleston, Dorchester, Berkeley, Georgetown, and Horry).

#### 2.4. METHODS: DATA COLLECTION AND PREPROCESSING

Our research approach is based on different datasets, study areas, and timeframes (Figure 2.2). Over 418 million geotagged tweets within the continental U.S. were collected using the Twitter Stream Application Programming Interface (API), comprising a six-month period from July 7th, 2016 until December 31th, 2016. These tweets are stored and managed in a Hadoop (<http://hadoop.apache.org>) environment that served as tweet

repository for this study. The repository was then queried with Apache Impala (<https://impala.incubator.apache.org>) using spatiotemporal criteria and keywords in the tweet message and hashtags, obtaining seven Twitter datasets covering different spatial regions, keywords, time periods, and spatial accuracies (Table 2.1). The locational accuracy of a geotagged tweet depends on how a Twitter user shares his/her location when posting a tweet. The location can be shared in the format of place names (e.g. country, state, city) or the exact latitude and longitude (point-level, determined by the device's GPS or other signals such as cell tower).

#### 2.4.1. DATASETS FOR SPATIOTEMPORAL ANALYSIS

The *Temporal Matthew-related* dataset contains the tweets (64,189) from the five affected states during the time period Matthew was active in the region (October 2nd±October 11th, 2016). Four keywords were applied in the query ("matthew", "hurricane\*", "evac\*", "storm\*") in accordance with other hurricane studies (Gupta et al., 2013; Kryvasheyev et al., 2016). In this case, the lowest spatial accuracy was kept to the state level. The wildcard "\*" was used for fuzzy query to include other variants of the keyword, such as hurricanes and "hurricanematthew" (hashtag).

The *Temporal Evacuation-related dataset* is a subset of the previous dataset composed of 7,735 tweets in the states that ordered voluntary or mandatory evacuations for Matthew (Florida, Georgia, and South Carolina). The timeframe and the spatial accuracy remained unaltered while the keywords of the query were reduced to "evac\*".

The spatial analysis requires a finer degree of accuracy than the temporal analytical set. Therefore, the *Spatial Matthew-related* dataset only includes the tweets with city-level and

point-level accuracy. The keywords were kept similar to the temporal analysis. The number of tweets contained in this dataset was 54,509.

To determine the universe of Twitter users in the study area, a dataset (*Twitter Population*) containing 4,706,685 tweets within the time span from September 15th, 2016 to October 15th, 2016 was extracted for the five affected states. The keyword was left blank to collect as many tweets as possible. The spatial accuracy needed matched with the requirements of the analysis (at least in city-level).

#### 2.4.2. DATASETS FOR EVACUATION BEHAVIOR ANALYSIS

South Carolina was used for the coastal evacuation analysis. Thus, the *Pre-evacuation* dataset identifies the largest possible number of users in the coastal counties prior to the evacuation. Consequently, the keyword of the query was left blank. The time span encompassed from October 2nd, 2016 to October 4th, 2016, with a spatial accuracy set in the city-level or higher. A total of 13,370 tweets formed the dataset after a manual filtering process to remove nonhuman (bots) users.

A *Post-evacuation* dataset was extracted for the time in which Hurricane Matthew was affecting South Carolina (October 7th 6pm, 2016 - October 9th 10am, 2016). We assumed that by October 7th at 6pm the evacuation had been completed, taking into consideration the night (Dow and Cutter, 2002) and travel (rainy) conditions (first effects of Matthew). In a similar manner, we understand the return period had not been initiated by October 9th at 10am considering that the trip would have implied driving at night and in rainy and flood conditions. We used the active users from the *Pre-evacuation* dataset in order to narrow down the query to the users that posted in both periods of time (pre and post evacuation).



The keyword was left blank (\*) in order to gather as many tweets as possible while the spatial accuracy was set on the city-level or higher, covering the United States. This dataset gathered 13,685 records.

The *Return* dataset was collected to analyze the return timeline of the evacuees. Therefore, it gathers the tweets posted from the end of the post-evacuation period until October 19th, 2016. Spatial accuracy remains in the city-level and the keyword was left blank. This dataset contained 19,216 tweets throughout the United States.

## 2.5. TWITTER REACTION TO HURRICANE MATTHEW

### 2.5.1. TEMPORAL ANALYSIS

To determine the temporal trends on Twitter activity related to Hurricane Matthew, we generated six hourly plots of Matthew-related tweets. Daily aggregates were also included in the plots to facilitate the overall interpretation.

In the affected states (Figure 2.3a), Hurricane Matthew-related tweets started increasing in the afternoon of October 3rd as a consequence of the strength of the hurricane, the impending landfalls in Haiti and Cuba, and the revised forecast of movement in a northern direction with a projected landfall somewhere between Florida and North Carolina. Altogether, the tweeting activity peaked on October 6th and then declined rapidly thereafter. This pattern is highly affected by the timing of the tweets from Florida, as they represent about 60% of the total of tweets in the study region. Regionally, as the forecasted track changed, there was a gradual shift in social media attention between October 6th and October 9th. Twitter activity in Florida peaked on October 6th and then declined the following day. Georgia also peaked on October 6th, but the drop on October 7th is less

noticeable. For North Carolina, the social media attention reached its highest on October 8th, the day the hurricane made landfall in South Carolina. Virginia, the northernmost state in the study area, still had a large volume of Twitter activity on October 9th (as the hurricane windfield was still affecting its coast).

South Carolina is an anomaly in Twitter trends. Twitter activity peaked on October 4th, dropped the next day, and then gradually reached a second peak on October 8th. This pattern reflects two events. The first peak was triggered by the official mandatory evacuation order given by the Governor South Carolina, which caused a blast of attention to the hurricane, as it served as official confirmation of its potential danger. The second peak was caused by the hurricane's proximity to South Carolina's coast and impending landfall.

To further investigate Twitter reaction to the evacuation orders, we plotted the evacuation-related tweets for three states that had evacuation orders issued: Florida, Georgia, and South Carolina. As illustrated in Figure 2.4, evacuation-related tweets showed a different temporal pattern between the states. The South Carolina evacuation order, on October 4th was pre-emptive and heightened Twitter attention in both South Carolina and Florida. Georgia and Florida experienced their peaks on October 6th, corresponding to the evacuation orders issued by the two states.

The overall temporal patterns identified in Figure 2.3 diverges from earlier studies of social media reactions to hurricane threats such as Irene in (Mandel et al., 2012) or Sandy in 2012 (Huang and Xiao, 2015; Kryvasheyeu et al., 2016; Murthy and Gross, 2017), where the peak of the response was reached when the landfall occurred. Matthew-related content

showed a different dynamic, as maximum Twitter activity occurred before the storm's most intense rainfall and winds were felt in the affected states. In fact, at the time of the landfall, Twitter activity in the study area had dropped considerably. This may have been caused by the unusual track of the storm –an offshore parallel path and the overall weakening of the storm after landfall. Also, this region of the US has more experience with hurricanes and evacuation, resulting in residents paying more attention to preparatory actions in advance of the storm, rather than waiting for landfall (Figure 2.4). The immediacy of the situational awareness provided by Twitter certainly enhances preparedness in advance and during the storm, which is why it is so widely used by residents and officials (Wukich, 2015). However, there is at least one case study that illustrates the utility of using Twitter during the restoration phase, as was the case with Typhoon Haiyan in 2013 (David et al., 2016).

### 2.5.2. SPATIAL ANALYSIS

The spatial analysis was carried out at the county level for the whole study area by counting the number of Matthew-related tweets per county with the spatial join function provided by ArcGIS. Following Shelton et al. (2014), we normalized the Twitter activity by the Twitter population (active Twitter users from September 15th to October 15th in the study area) to reduce the population bias. The ratio between the number of Matthew-related tweets and the Twitter population accounted for the normalized Matthew-related Twitter Activity (MTA) for each county.

$$MTA = \frac{\text{Matthew – related Tweets}}{\text{Twitter Population}}$$

To obtain the active users within the counties of the affected area (Twitter population), we used the Twitter population dataset (Table 2.1). We assigned the location where a user

tweeted the most during the time period as their "home" location, using only those users with more than ten tweets in the same city to ascertain that the sample represents the active local users. This threshold of 10 tweets was established based on the distribution of the tweets per user. To avoid outliers in Twitter activity caused by the small amount of users, only those counties with more than 50 active users (151 counties) during that timeframe were used for further analysis. It is worth noting that the subset of 151 counties contains 96.7% of the Matthew-related tweets, as it includes the most populated counties in the study area. Figure 2.5 shows the spatial distribution of the Matthew-related Twitter activity (MTA) for the subset of counties.

As illustrated in Figure 2.5, the most intensive Twitter activity was generally recorded along the Atlantic coast of the affected states. Florida exhibits more activity compared to the other four states. This is due to a number of factors: the size of the coastal population in the state, the length of its coastline, the coast-to-coast width of the state where residents on the western coast could be affected, and the decreasing magnitude (category) of the hurricane as it traveled northward. These factors coupled with initial intensification and the uncertainty of the forecasts combined to produce a greater social media response in Florida, even in counties with no Atlantic coastline.

To further explore the spatial relationship between Twitter activity and Matthew, we conducted a regression analysis between the normalized Twitter activity (MTA) per county and the distance from the county centroid to the estimated track of the hurricane ( $d$ ) (sample size: 151). As shown in Figure 2.6, Twitter activity and the distance to hurricane exhibit significant negative correlation, with  $r$  of -0.83. The negative relationship is illustrated by a power best-fit line, showing a sharp decline in the normalized number of tweets as

distance increases. This result confirmed the visual pattern of the map (Figure 2.5). More importantly, the best fit line and its mathematical model (the equation in Figure 2.6) is consistent with the well-known Distance-Decay function (inverse power law format with the exponent  $\beta = 1.26$ ), which describes the effect of distance on spatial and social interactions.

The spatial analysis illustrated that the closer a community is to the threat, the more likely it is to engage in related social media content. This finding confirms and reinforces studies previously conducted for other disaster events (de Albuquerque et al., 2015; Kryvasheyev et al., 2016; Li et al., 2018b). This trend may be driven by the likelihood of personally experiencing the effects of the storm, the individual perception of risk, and the social media discussion about the preparation actions.

## 2.6. RESPONSE TO THE EVACUATION ORDER

In the evacuation behavior analysis, we downscaled our study area to three South Carolina Coastal Hurricane Conglomerates used for planning and response actions. These contain eight coastal counties: Horry and Georgetown (Northern Conglomerate); Charleston, Berkeley, and Dorchester (Central Conglomerate); and Beaufort, Colleton, and Jasper (Southern Conglomerate). The focus was on a more detailed analysis of actual evacuation behavior at this local scale –how many Twitter users evacuated, their evacuation destination, and return date.

We first identified the local active Twitter users during the pre-evacuation period (October 2nd to 4th) based on the following steps (Figure 2.7): 1) extraction of the Twitter users who posted at least one message during the pre-evacuation and the post-evacuation

timeframe (pre-evacuation and post-evacuation dataset, Table 2.1); 2) extraction of all tweets (city-level, point-level) posted by the users obtained in the previous step from the Geotagged Tweets Repository (Table 2.1). To limit leisure trips, the national holiday periods (Thanksgiving, November 22-29 and Christmas, December 20-31) were excluded; 3) computation of the median center of the location (coordinates latitude and longitude) of all tweets posted by each user using the median center function provided by ArcGIS. If the location accuracy of a tweet is city-level, the centroid of the city boundary was used as the coordinates for the calculation; and 4) The median centers were then considered as the "home" location for the users, therefore separating the 1,384 local users (median center within the coastal counties, 79.3%) from the 361 transient visitors (20.7%). Lastly, we calculated the median center of the tweets posted by the 1,384 users during the post-evacuation period (when Matthew was directly affecting South Carolina). This provided a post-evacuation location for each user.

#### 2.6.1. HOW MANY USERS EVACUATED?

Figure 2.8a shows the home location of the 1,384 active local users identified before the evacuation, especially clustering around two major cities: Charleston (Central Conglomerate) and Myrtle Beach (Northern Conglomerate). Figure 2.8b displays the locations of the same 1,384 users during the post-evacuation period. Blue dots represent the users who moved outside of the eight coastal counties, and red dots show the location of the users who did not evacuate. Table 2.2 depicts the number of evacuated users in the three conglomerates and the estimated number of evacuees based on the population of each conglomerate. We only considered users as evacuees if they posted all their tweets from beyond the coastal counties during the post-evacuation period. Thus, 37 users (2.7%) were

classified as not evacuated, since they tweeted from both within and beyond the coastal counties during our post-evacuation time period.

For the entire study area, 54.0% (747 out of 1,384) of the local users evacuated, a finding consistent with previous evacuation studies. For example, the South Carolina hurricane evacuation behavioral study, Cutter et al. (2011) indicated that 76.6% of respondents would evacuate for a major hurricane while only 21% expressed their intention of evacuating for a weaker hurricane. Our results from an actual evacuation fall in a middle position between the two intentions based on storm category. Although Matthew finally affected the state as a Category 1 hurricane, at the time the evacuation order was given the forecasted intensity was as a Category 3 (Figure 2.6). This may explain the higher evacuation compliance observed in our study in comparison with Cutter et al. (2011), as the decision of evacuating is made upon forecasts and not storm intensity at landfall. In a fairly similar hurricane in terms of intensity, Hurricane Floyd in 1999, Dow and Cutter (2000) reported an actual evacuation rate of 65%, which also supports our evacuation estimation. Another interesting finding that also lines up with Cutter et al. (2011) is the noticeable geographic variation among the three conglomerates. The Southern Conglomerate experienced the highest evacuation rate (77.1%), followed by Northern Conglomerate (53.7%) and Central Conglomerate (50.0%). The South Carolina behavioral study (2011) also identified more willingness to leave in the Southern Conglomerate (85.7% in major hurricanes and 32% in weaker hurricanes) than in the Central Conglomerate (74.2% - 18.0%) and Northern Conglomerate (70.7% - 12.0%). With nearly 1.36 million people living along South Carolina's coast in 2015, we estimated that 760,000 people evacuated in response to Hurricane Matthew. Based on South Carolina daily traffic

counts, approximately 355,000 vehicles left the coastal counties during the period October 4-7th (SCDOT, 2017). As reviewed previously, most traffic-oriented studies only estimate the number of vehicles or the number of vehicles per household taken, not the number of people in each car (Cutter et al., 2011; Archibald and McNeil, 2012). While the average household size in South Carolina coastal counties ranges from 2.42 to 2.84, many of these households live outside of mandatory evacuation zones. Rather than take this aggregate figure of household size as our occupancy per vehicle during the evacuation, we assumed a more conservative average occupancy of 2.0 people per car estimating that 710,000 people likely evacuated from the coastal counties –a number consistent with our estimates based on Twitter.

#### 2.6.2. WHERE DID RESIDENTS TRAVEL TO?

Most of the Twitter users were concentrated near the population centers of Charleston and Myrtle Beach. To visually represent destinations of the evacuees, locations were aggregated to the state scale (Figure 2.9). The majority of evacuated Twitter users (45.6%) did not leave South Carolina (Table 2.3). For those who moved out of state, 18.3% traveled to North Carolina and 9.1% to Georgia in areas far away from the coast. This confirms people's propensity to evacuate to places relatively close to their homes, based on proximity to family and friends and the availability of hotels (Cutter et al., 2011). Also worth mentioning is the significant percentage of population evacuating to northeastern states (15.0%) such as Virginia, Pennsylvania, Maryland, New Jersey, and New York. In the study conducted by Cutter et al. (2011), up to 70% residents reported that their potential evacuation shelter options were a friend's or relative's home (40.7%) or a hotel/motel



(29.9%). The presence of family/friends, a larger accommodation capacity, and a good transportation corridor (I-95) may account for the northeastern destination preferences.

Surveys of South Carolina residents who evacuated from hurricanes Bertha and Fran in 1996 revealed that 15% and 28%, respectively, evacuated out-of-state (Hazards Management Group Inc 1985; Baker 1997). During Hurricane Floyd, 56% of South Carolinians who evacuated left the state (Dow and Cutter, 2002). This last figure matches closely with the 54.4% registered in our assessment, although it is considerably higher than the 35% reported by Cutter et al. (2011) for a hypothetical evacuation.

### 2.6.3. WHEN DID EVACUEES RETURN?

To investigate the return behavior, we tracked the locations of each tweet from the 747 evacuated users on a daily basis from October 9th to October 19th (Return dataset, Table 2.1). If the user's Tweet location was within the eight coastal counties during this time period, then that user was counted as returned evacuee. People, against official recommendation, started to return home on October 9th (11.0% of evacuated users returned) (Table 2.4). The following four days (October 10th - 13th) experienced the highest returns rates (18.6%, 14.9%, 15.1%, and 11.9%, respectively). October 14th registered a noticeable drop in the return rate, which gradually slowed down from 4.8% to 2.3% on October 10th. As of October 19th, 88.0% of the users had returned while 4.0% remained away from their home locations. The remaining 8.0 percent had not posted any messages during the 11-day period, therefore making it impossible to identify their locations.

There are no prior studies of a timeline on the re-entry for South Carolina so we are unable to compare our results to previous hurricane experiences in the state. It is also possible that the data underestimate the return date for some evacuees simply based on no Twitter posts from individuals. This is an avenue for further research.

## 2.7. LIMITATIONS

There are a number of limitations to the approach used in this paper. First and foremost is the representativeness of Twitter data, which may not reflect the characteristics of the population under examination in terms of gender, race/ethnicity, socio-economic status, or age. While the emergent literature provides several attempts to understand the demographic profile of Twitter users (Mislove et al., 2011; Hecht and Stephens, 2014), samples of Twitter users do have urban, gender (male), and race (Caucasian) biases. In other words, rural areas are less represented, as are females, and non-white users. However, the representativeness shortcoming is also shared with questionnaire surveys where respondents are often older, more educated, and include fewer minority respondents than the demographics of the area would suggest. The main difference between the two samples is that in the latter the researcher knows the biases (based on self-reports of respondents) and weights the results accordingly. The lack of any personal information about Twitter use (other than location) precludes knowing the representativeness of the sample.

Another limitation regarding the representativeness of the sample is the long-tail effect of the social media contributions where most of the social media contents (e.g. tweets) are contributed by a few users (Li et al., 2013). One implication of this effect is that active users can represent large numbers of individuals. For example, a single Twitter user who posts 25 tweets a day may weigh the same as 25 Twitter users who post once a day. In this

sense, the spatiotemporal patterns of Twitter activity identified in the paper do not represent the total population, only accounting for the reaction of social media. This representation issue also applies to the evacuation analysis because the active users have a higher chance to be selected as the evacuation sample (1,384 users) based on the steps described in Section 6. More research is needed to improve the capacity to infer the representativeness of Twitter samples regarding both demography and total population.

A third limitation is the selection bias of the available social media data, as people select to be included or to share their data, rather than statistically sampled. For example, Twitter users must grant consent to offer the geolocation of their tweets, so the locational information (origin - destination) could be highly skewed. It is unclear, for example, exactly what percentage of total Tweets contain geotags and how this might vary geographically during an emergency or disaster situation. In addition, since the Twitter streaming API only provides a small portion of all posted tweets, it is unclear how the streamed geotagged tweets were sampled from the tweets population.

Lastly, our approach is unable to systematically ascribe specific factors or motivations for the decision making to evacuate including the destination choice. The evacuation decision making during a hurricane is quite complex, requiring to consider different dimensions of evacuee behavior, such as the route choice, evacuation mode, choice of safe destination or mobilization time. There are extensive existing studies along this line and particularly notable are the work by Sadri et al. (2013, 2014a, 2014b, 2015) and Hasan et al. (2010, 2013a). Since this paper is not about evacuation decision making, but merely documenting the overt behavior (departure time, re-entry time, and location) as evidenced by social media data (tweets), the discussions of the underlying motivations of the route

choice, choice of destination, and return time are beyond the scope this paper. The scope is also constrained by the limited information we can retrieve from the tweets comparing to traditional questionnaire surveys. However, since social media data contain social network information (e.g. followers, friends), we see a potential research avenue of utilizing social media data for understanding evacuation decision making, complementing the existing approaches of using questionnaire surveys or interviews (Sadri et al., 2017a; Sadri et al., 2017b).

## 2.8. CONCLUSIONS

This article examined Twitter reaction both spatially and temporally in response to a hurricane as a potential innovative approach for assessing protective action behavior and evacuation compliance in a timelier and cost-efficient manner. Considering that nearly half of the state level emergency management agencies in the United States intend to leverage social media and open source information for public and situational awareness (San et al., 2013) this research confirms the utility of social media in monitoring public awareness and evacuation behavior in response to Hurricane Matthew.

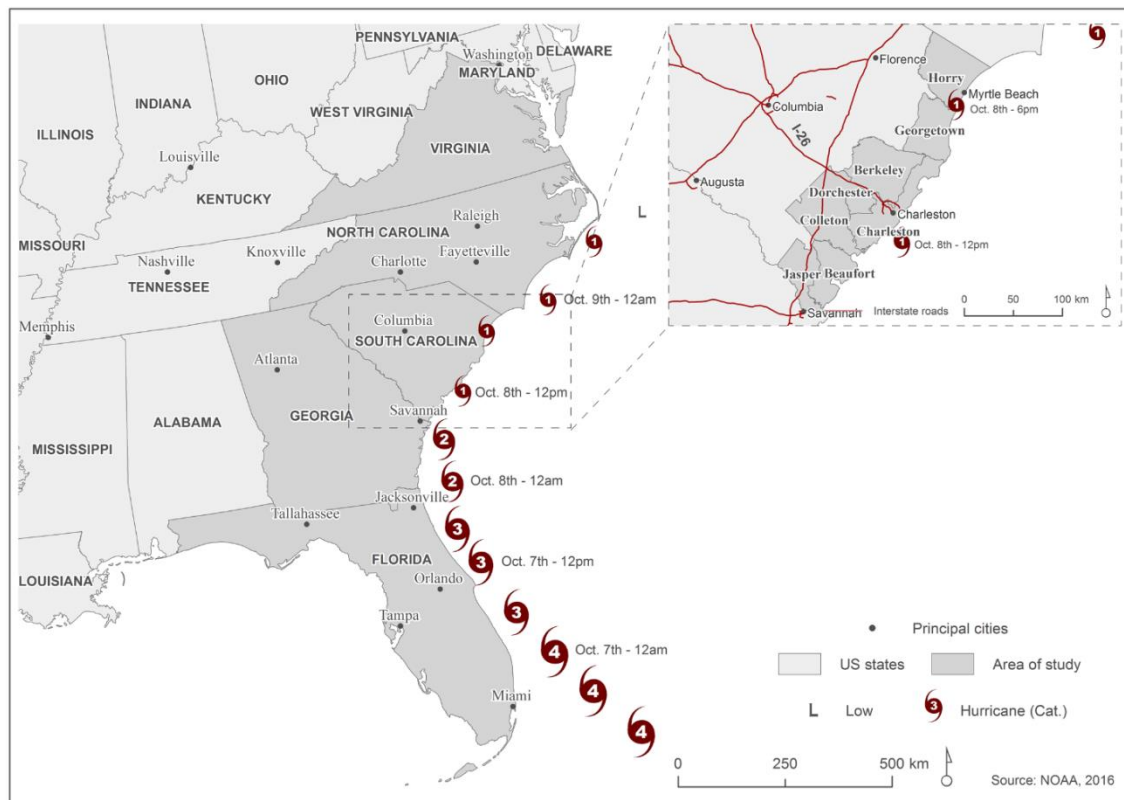
Conducting spatiotemporal analyses permits the examination of public behavior as the hurricane moved northward. It also allows us to gauge the differential response by state emergency managers to the perceived threat of the hurricane based on their preparedness activities and calls for mandatory evacuations. Only Governors of the affected states can issue mandatory evacuation orders and the timing and extent of such emergency actions varies considerably even among neighboring states. In this sense, we observed a different temporal pattern in Twitter response during a tropical cyclone to the ones previously identified in the literature, as the most active period was recorded in advance to the advent

of the storm and was linked to actions in the preparedness stage and distance from the storm track. A more detailed analysis of South Carolina evacuees using pre-evacuation and post-evacuation Twitter geotagged data enabled an estimation of the compliance rate for the mandatory evacuation, the timing of evacuee departure and return, and the destination of evacuees seeking shelter from the storm.

Our approach offers a solution to tackle several of the drawbacks of traditional assessments of evacuation rates through questionnaire surveys. These are frequently time-consuming and costly, and the response rates are often far from ideal. This alternative is cost-efficient and timely as the data can be collected in real-time, providing a remarkable sample size with successful results, as it has been shown. While the approach does have its limitations as mentioned previously, we do have a more robust measure of when Twitter users likely left, returned, and where they went (e.g. destination county). So the trade-off is the immediacy and relative accuracy of the timing of the evacuation and the destination of the evacuees versus the detailed motivation behind such behavioral responses assessed months later.

We believe the advantages outweigh the shortcomings of this approach to monitoring evacuation behavior. It provides complementary and near-real time data for assessing evacuation responses, and can be very useful when examined in tandem with traditional evacuation behavior methods. Future work should test the suitability of this approach in other emergency situations, as well as investigating the representativeness of Twitter samples for evacuation studies. Another line of research may take advantage of other sources of social media data. In this sense, it is worth noting that the pool of data available is considerably larger than what was used in this study (consider Facebook -79% of online

adults-, Instagram-32%-, Pinterest-31%- in comparison with Twitter -24%-, (Greenwood et al., 2016)). A third line of research is to analyze the network properties and tweeting activities (Sadri et al., 2017c) of the 1,384 Twitter users to identify the potential social factors that influence their decision to evacuate or stay, with an aim to predicting what kind of users are likely to evacuate via social network amplification (Kryvasheyeu et al., 2016). Finally, research on the integration of new and traditional approaches and data sources is needed to explore avenues for better understanding evacuation decision making and improving evacuation management.



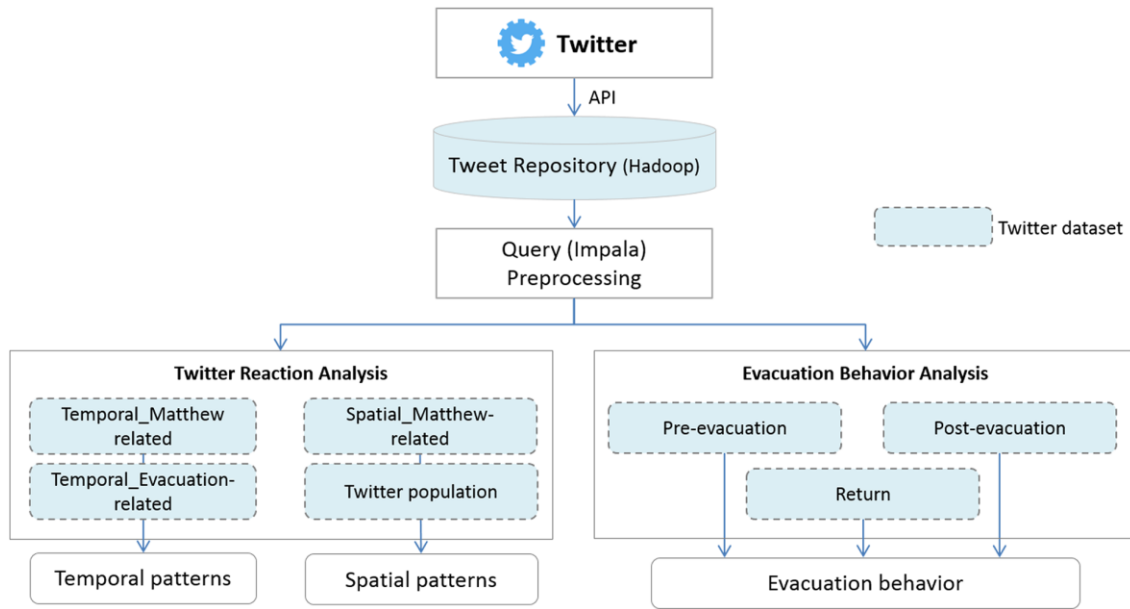


Figure 2.2. Overview of the research approach and the Twitter datasets used



Table 2.1. Twitter datasets

	<b>Dataset name</b>	<b>Region</b>	<b>Keywords</b>	<b>Time period</b>	<b>Location accuracy</b>	<b>Number of tweets</b>
Temporal analysis	Temporal Matthew-related	FL, GA, SC, NC, VA	matthew, hurricane*, evac*, storm*	10/02 - 10/11	State, City, Point	64,189
	Temporal Evacuation-related	FL, GA, SC	evac*	10/02 - 10/11	State, City, Point	7,735
Spatial analysis	Spatial Matthew-related	FL, GA, SC, NC, VA	matthew, hurricane*, evac*, storm*	10/02 - 10/11	City, Point	54,509
	Twitter Population	FL, GA, SC, NC, VA	*	09/15 – 10/15	City, Point	4,706,685
Evacuation behavior	Pre-evacuation	SC coastal counties	*	10/02 - 10/04	City, Point	13,370
	Post-evacuation	US	*	10/07, 6pm - 10/09, 10am	City, Point	13,685
	Return	US	*	10/09, 10am – 10/19	City, Point	19,216

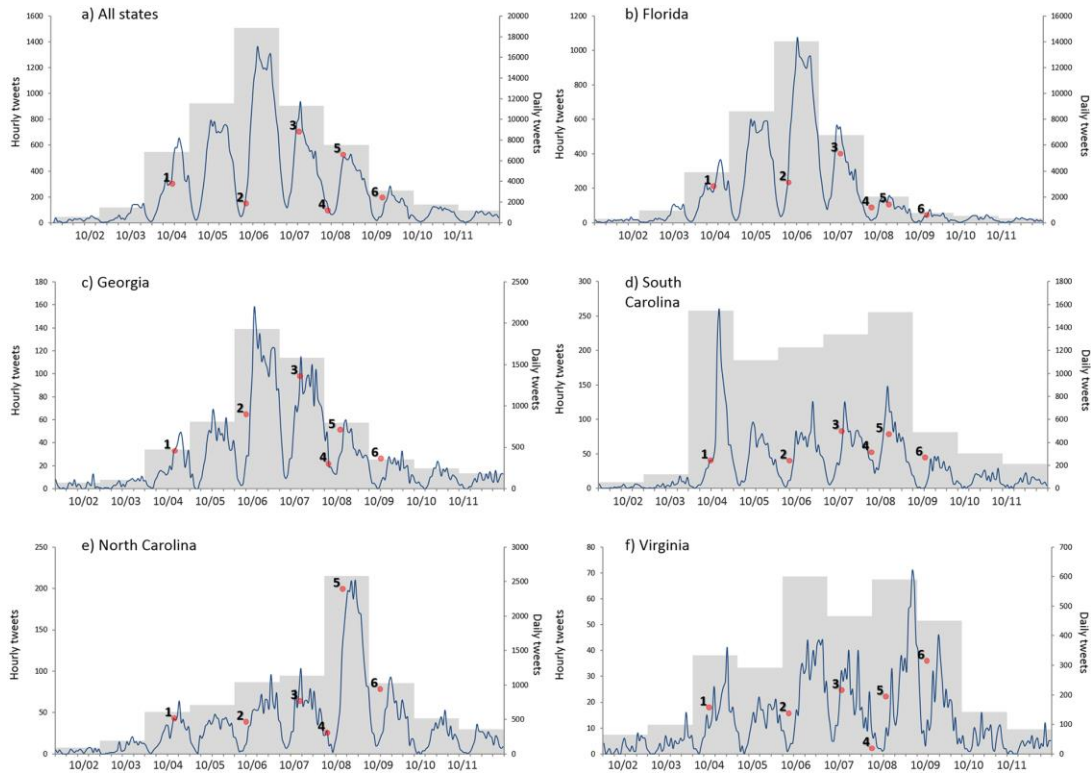


Figure 2.3. Temporal distribution of Matthew-related tweets for the five states. a) All states b) Florida, 59.7% of the tweets c) Georgia, 10.2% of the tweets d) South Carolina, 12.4% of the tweets e) North Carolina, 12.9% of the tweets f) Virginia, 4.9% of the tweets. Shaded area represents daily tweets, while the hourly tweets are shown on the line graph. Numbered dots indicate the selected key events during the time period: 1. South Carolina evacuation order. 2. Georgia evacuation order. 3. Hurricane eye offshore Daytona Beach (FL). 4. Hurricane eye offshore Savannah (GA). 5. Hurricane eye offshore Charleston (SC). 6. Hurricane eye offshore Wilmington (NC).

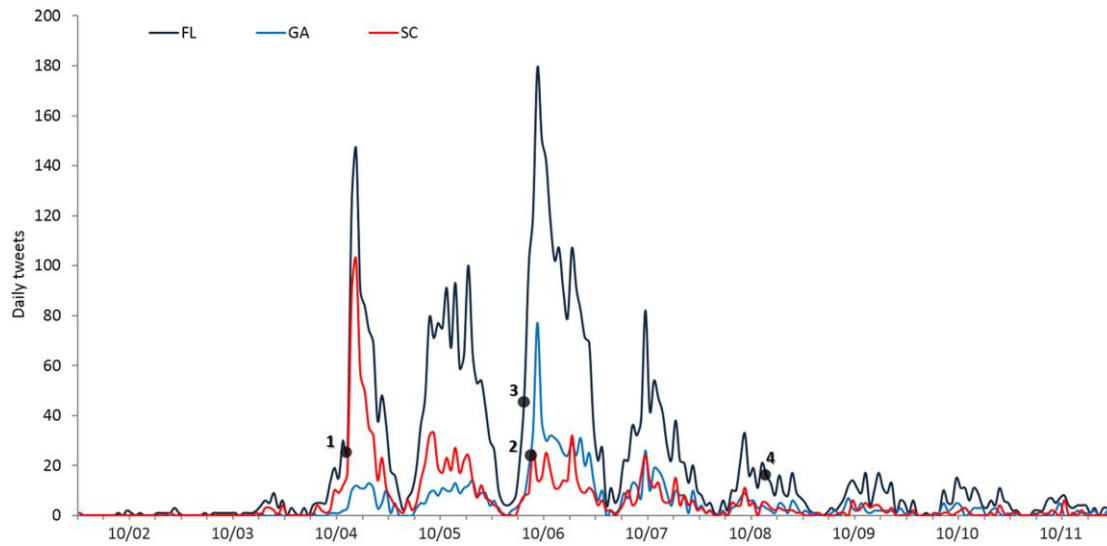


Figure 2.4. Temporal distribution of evacuation-related tweets. Numbered dots indicate the key events during the time period: 1. South Carolina evacuation order (SC residents in the coastline counties were mandated to evacuate on October 4th); 2. Georgia evacuation order (GA residents on the east of Interstate 95 were mandated to evacuate on October 6th); 3. Florida evacuation warning by Florida Governor (1.5 million of residents living in evacuation zones on the Atlantic coast of Florida were urged to evacuated on October 6th); 4. Landfall near McClellanville, South Carolina (October 8th, 1500UTM).

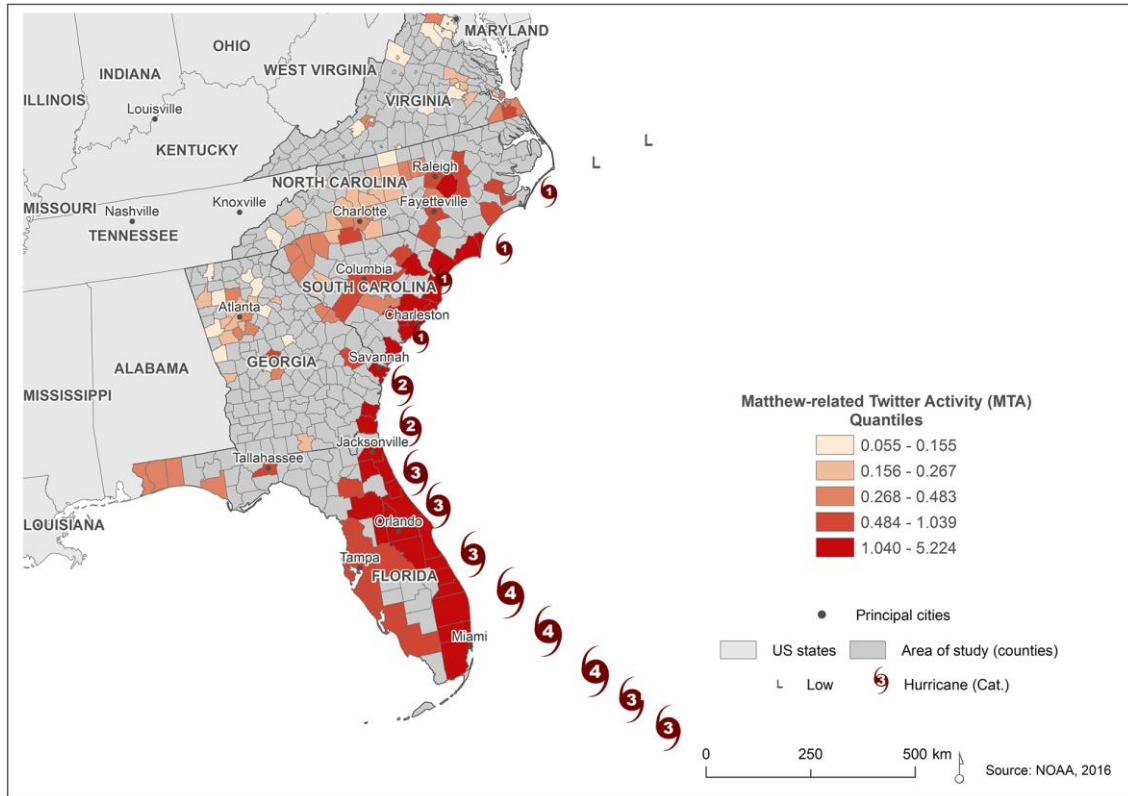


Figure 2.5. Spatial distribution of Matthew-related Twitter activity in counties with more than 50 active users.

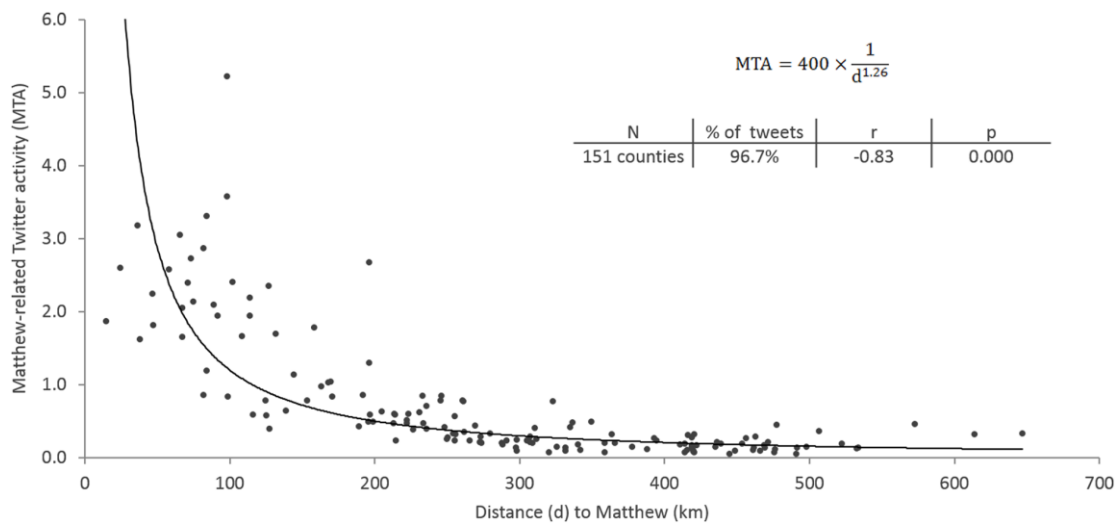


Figure 2.6. Correlation between Twitter activity per county and distance to Hurricane Matthew.

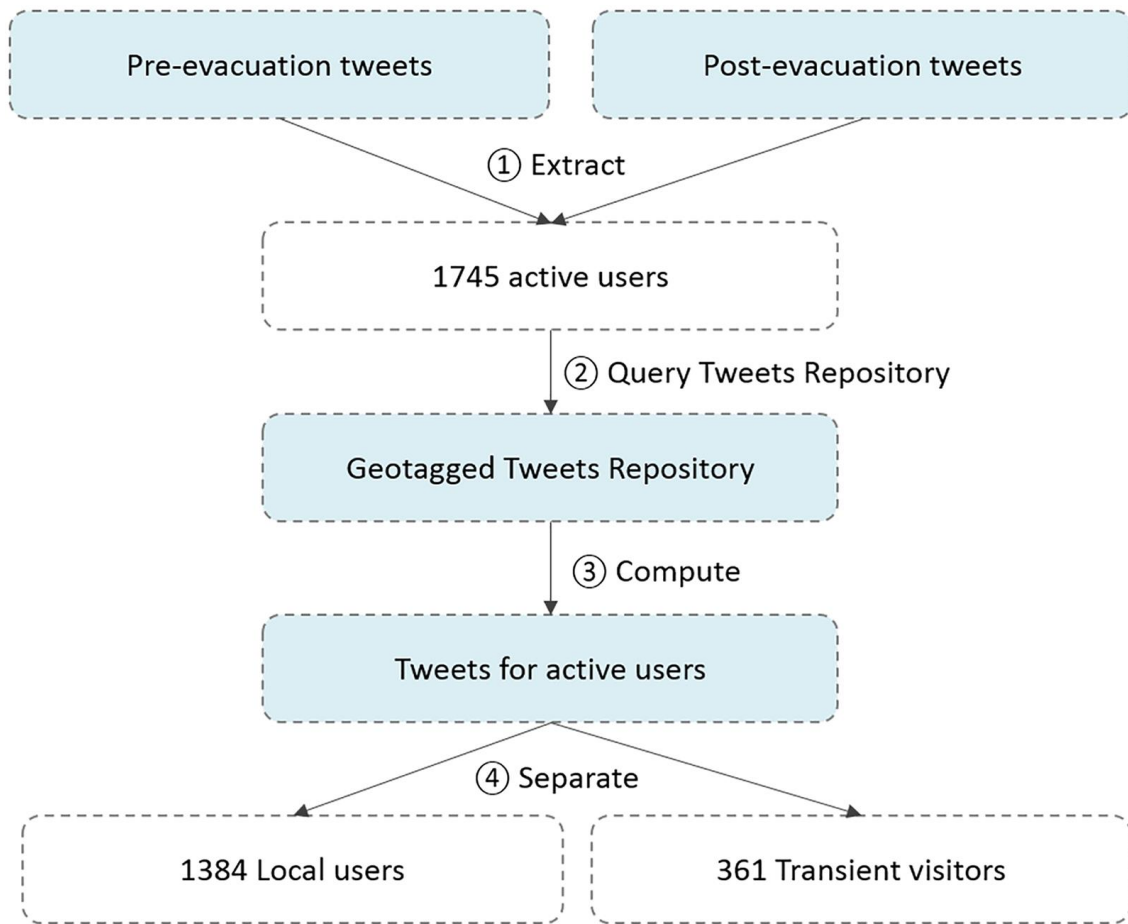


Figure 2.7. Workflow to obtain the local users.

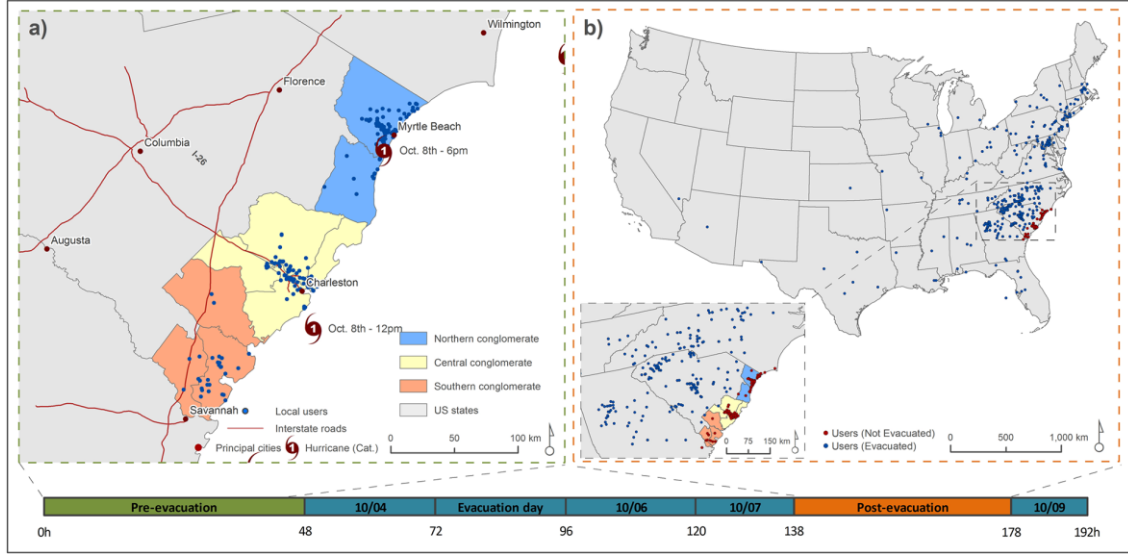


Figure 2.8. Response to the evacuation order of the 1384 local Twitter users. (a) Local-user locations during pre-evacuation period (10/02 - 10/04). (b) Local-user locations during post-evacuation period (10/07 6pm - 10/09 10am). Blue dots represent the users who moved outside of the risk area (eight coastal counties), and red dots were the users who did not evacuate.

Table 2.2. Evacuation rates and estimated evacuees by conglomerate

Conglomerate	Users	Evacuated	% Evacuated	Pop. 2015	Estimated evacuees
Southern	105	81	77.1%	245,144	189,111
Central	560	280	50.0%	744,526	372,263
Northern	719	386	53.7%	370,497	198,904
Total	1,384	747	54.0%	1,360,167	760,278

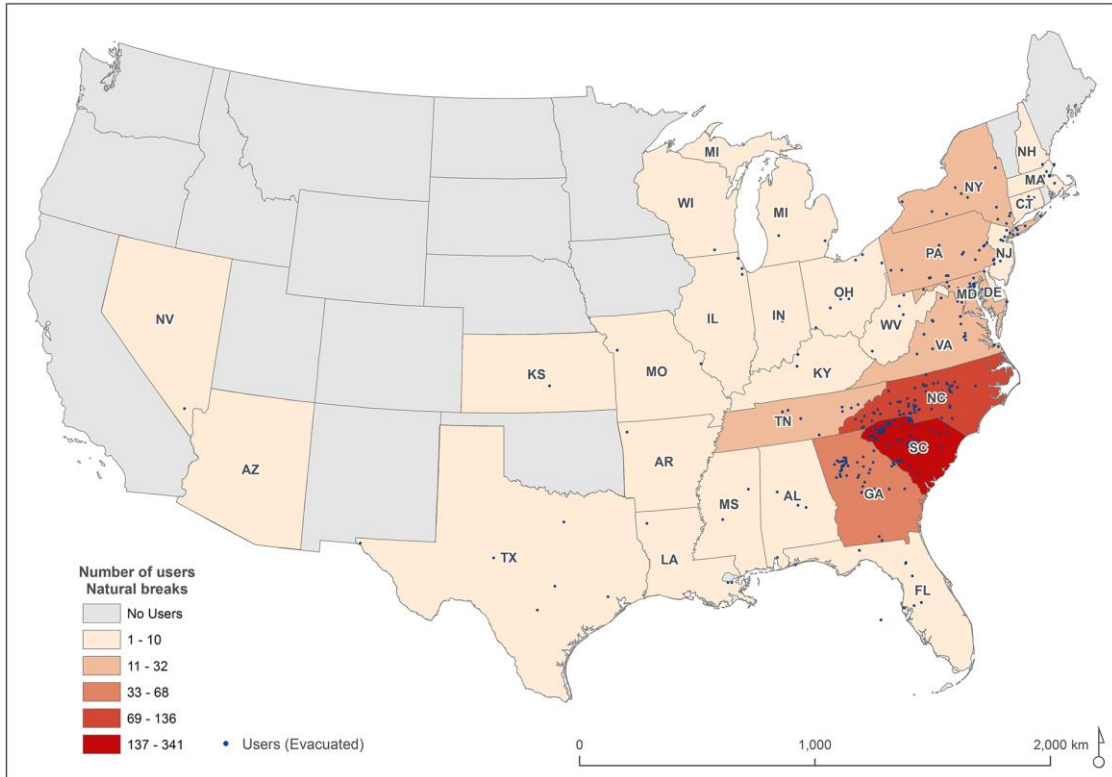


Figure 2.9. Evacuation destinations from South Carolina during Hurricane Matthew.

Table 2.3. Evacuation destinations from South Carolina during Hurricane Matthew

State	Users	%
South Carolina	341	45.6%
North Carolina	136	18.2%
Georgia	68	9.1%
Virginia	32	4.3%
Pennsylvania	26	3.5%
Maryland	24	3.2%
New York	20	2.7%
Tennessee	19	2.5%
New Jersey	10	1.3%
Other states	71	9.5%



Table 2.4. Return after evacuation

<b>Date</b>	<b>Users</b>	<b>% Returned</b>	<b>% Total</b>
Sun. 10/09 from 10am	82	11.0%	11.0%
Mon. 10/10	139	18.6%	29.6%
Tue. 10/11	111	14.9%	44.4%
Wed. 10/12	113	15.1%	59.6%
Thu. 10/13	89	11.9%	71.5%
Fri. 10/14	36	4.8%	76.3%
Sat. 10/15	27	3.6%	79.9%
Sun. 10/16	19	2.5%	82.5%
Mon. 10/17	18	2.4%	84.9%
Tue. 10/18	17	2.3%	87.1%
Mon. 10/19	17	0.8%	88.0%
Not tweeted 10/09-10/19	60	8.0%	
Not returned	30	4.0%	
<b>TOTAL</b>	<b>747</b>	<b>100.0%</b>	

## **CHAPTER 3**

### **BRIDGING TWITTER AND SURVEY DATA FOR THE EVACUATION ASSESSMENT OF HURRICANE MATTHEW AND HURRICANE IRMA<sup>2</sup>**

---

<sup>2</sup> Martín, Y., Cutter, S.L., & Li, Z. Bridging Twitter and survey data for the evacuation assessment of Hurricane Matthew and Hurricane Irma. Submitted to *Natural Hazards Review*

### 3.1 INTRODUCTION

Tropical cyclones are a major destructive and disruptive agent in the life of hundreds of millions living within the risk area for such phenomena. Severe winds, storm surge, heavy rainfall, and tornadoes bring death, property damage, and economic losses to multiple regions in any given year (Hsiang and Jina, 2014). During the last hurricane season (2017) in the Atlantic basin, one of the most active on record (NOAA, 2018), these events extended havoc across many countries, causing thousands of direct and indirect deaths and millions of displaced residents. Fortunately, meteorologists have developed monitoring and predictive systems that allow the issuance of hazardous tropical weather watches and warnings several days in advance. Ultimately, this offers the communities potentially at risk a small window to initiate protective actions to minimize the loss of lives and mitigate property damage. In this sense, evacuations are the primary protective strategy taken by those who live near the shoreline, where storm surge is the biggest concern (Cutter and Smith, 2009).

Efforts by researchers coming from multiple backgrounds (geography, psychology, communication, mobility, or civil engineering) have advanced the understanding of evacuation decision-making of people at risk, identifying some of the most common drivers and patterns of evacuation behavior. Most of this knowledge acquired through the use of survey questionnaires permitted gauging the response and motivation of different demographic groups to evacuation orders. However, some authors reported weaknesses in the method and claimed that some surveying approaches fail to efficiently provide essential information demanded by emergency managers, leading Baker (2009, p. 209) to comment: more creative methods or combinations of data-collection methods need to be employed.

In the age of technology and communication, leveraging Information and Communication Technology (ICT) and Big Data and associated analytics in disaster risk reduction is one of the most promising opportunities to reduce the impact of hazards (Data-Pop Alliance, 2015). For example, during the response phase of emergency management, Chen et al. (2016) was one of the firsts to point out the potential of social media in a hypothetical mass evacuation situation. Martín et al. (2017) tested the validity of Twitter for quantifying and characterizing a mass evacuation during a hurricane, reporting promising results.

This study expands the Martín et al. (2017) approach by relating the evacuation behavior of residents and non-residents and their demographic characteristics. Particularly, it investigates whether the evacuation behavior of racial minorities differs from racial majorities, whether young individuals are more prone to evacuate than middle-age or the elderly, and whether non-permanent residents find their way out of harm's way and in what proportion compared to local residents. Additionally, it explores the integration of traditional —survey— and innovative —Twitter— approaches and allows confirming the robustness of Martín et al. (2017) method for studying hurricane evacuation behavior.

## 3.2 RESEARCH CONTEXT AND BACKGROUND

### 3.2.1. THE SCIENCE: EVACUATION BEHAVIOR

Temporarily moving out of harm's way is one of the most effective protective strategies to cope with the impact of hazards. Evacuation decision-making and the behavioral outcomes are as varied as the specific initiating event and the affected population. Perceived risk, uncertainty, and the capacity to evacuate are the major drivers on the decision of evacuation or staying. Uncertainty, characterized by incomplete knowledge, is

distinguishable from perceived risk, an evaluation of uncertain events and impacts. Both influence the process of decision-making which might not appear to be logical or rational to outsiders (Golledge and Stimson, 1997). In fact, researchers identify risk perception as one of the leading factors in evacuation response (e.g. Dash and Gladwin, 2007; Huang et al., 2016a). Multiple predictors with theoretical influence over risk perception and the ability to evacuate have been extensively studied (Huang et al., 2016a) and include, risk area and distance (Zhang et al., 2004; Arlikatti et al., 2006); housing characteristics (Whitehead et al., 2000); warnings and storm information (Van Willigen et al., 2005; Morss et al., 2016); hurricane experience (Brommer and Senkbeil, 2010; Demuth et al., 2016); social networks (Sadri et al., 2017) and car ownership (Lindell et al., 2011), the latter an indicator of capacity to evacuate. Demographic characteristics that define different subpopulations have been investigated as well (Bowser, 2013; Dixon et al., 2017).

According to the literature, there are still gaps in our knowledge of how people react to threats (Bowser and Cutter, 2015) and inconsistencies in the empirical findings based on reviews and statistical meta-analyses (Baker, 1991; Huang et al., 2016a). For instance, few studies have investigated the connection between the length of residence (long-term and short-term residents) and the spatiotemporal response to hazards. Particularly, the response of short-term residents is not well understood (Phillips and Morrow, 2007) and is generally missing in traditional studies of evacuation behavior (Baker, 2009). Although some authors have specifically examined the evacuation behavior of tourists (e.g. Drabek, 1996; Cahyanto et al., 2014), comparative studies (same place, same event) are absent in the literature. The importance of understanding this short-term resident's spatial behavior increases as temporality becomes more and more relevant (Bauman, 1996) and people

become more mobile in both recreational and occupational endeavors (Hall, 2005). Specifically, Matyas et al. (2011) stressed the importance of considering the tourist population in evacuation research as they may lack knowledge about hurricane risks, be unfamiliar with their surroundings, and lack the support network of their local community or home. Other specifically under-studied groups in evacuation assessments are racial or ethnic minorities (Wilson and Tiefenbacher, 2012). For example, racial/ethnic minorities are less likely to prepare for emergencies (Burke et al., 2012) and ethnicity does have a significant effect on evacuation decision-making (Gladwin and Peacock, 1997; Huang et al., 2016b). However, a meta-analysis of hurricane evacuation studies reported non-significant effects of race (White, Black) and Hispanic ethnicity. Another group lacking further studying is the youngest segment of the adult population (Van Willigen et al., 2005), as traditional approaches —surveys— tend to be highly biased towards older respondents.

Obtaining a more comprehensive picture of the evacuation behavior of the overall population is a necessary input to advance in the adoption of self-protective strategies that minimize the vulnerability of people at risk. We believe an approach based on Twitter might open the possibility to study these three under-researched subpopulations —short term residents, racial and ethnic minorities, and the young— at larger scales than done previously.

### 3.2.2. DATA-COLLECTION METHODS

#### 3.2.2.1 TRADITIONAL APPROACHES

Classic methods for examining the mobility of populations in hazardous situations have mainly involved the use of questionnaire surveys and interviews. In the hazards and disaster field, survey questionnaires are the preferred data collection instrument for

studying hurricane evacuation behavior (Lindell et al., 2005; Siebeneck and Cova, 2012; Sadri et al., 2017). Researchers also utilize survey questionnaires to learn about the spatial behavior of residents to other hazards and at different scales such as tornadoes (Durage et al., 2014), earthquakes (Tamima and Chouinard, 2016), tsunamis (San Carlos Arce et al., 2017) or non-natural hazards such as nuclear accidents (Zeigler et al., 1981) and hazardous-material accidents (Mitchell et al., 2007).

Several features of surveys make them attractive for acquiring a significant level of knowledge of the driving factors determining spatial decisions (e.g. perception of risk, severity of the event, previous experience and family context) and planning the evacuation of large populations when danger arrives. For instance, the control that the researcher has over the sample and the questionnaire itself make surveys a flexible system to evaluate people's behavior, which is the most important advantage of this method. For evacuation assessment purposes, survey questionnaires offer a relatively high spatial and temporal resolution, as the respondent is often able to recall their movements and their timing, although some have expressed concern about a possible recall bias (Shareck et al., 2013). The richness of the background information of the respondent and motivation is also an important strength of this technique (Baker, 2009). On the other hand, surveys present several limitations that restrict its applicability as a unique and comprehensive method. First and foremost is the concern about the representativeness of the sample (Arlkatti et al., 2006). In recent decades, survey response rates have declined (Johnson and Wislar, 2012) and have often suffered from biases as respondents tend to be older than the overall population and minority groups are often under-represented (Bowser, 2013). In addition, surveys require extended collection times and resource demanding in terms of financial

and human investments (Baker, 2009). Lastly, surveys have difficulty reaching some sectors of population such as non-permanent residents (tourists or seasonal workers), and as a result provide insufficient information to characterize the evacuation behavior of the entire population at risk (Smith and McCarty, 2009).

The other method used for assessing spatial behavior in evacuations quantifies compliance with evacuation orders using Departments of Transportation traffic counters on egress routes out of mandated evacuation zones. A dense network of traffic gauges offers an accurate estimation of the number of cars on the road at every moment (real time data), which is a very valuable input for evacuation traffic management (Wolshon and Pande, 2016). Data accessibility and a low cost are some of the most regarded characteristics of this method (Ballard et al., 2008; Archibald and McNeil, 2012). A weakness in this data approach is the lack of information about the occupants of the vehicles. The researcher is unaware of the number of people per vehicle as well as their sociodemographic characteristics and the motivation behind their movement (Archibald and McNeil, 2012).

#### 3.2.2.2. PASSIVE CITIZEN-SENSOR DATA

Many studies have begun to exploit alternative methods to get a sense of people's spatial behavior under normal conditions (e.g. Chaintreau et al., 2007; González et al., 2008; Jurdak et al., 2015), but there is a paucity of studies under hazardous or disastrous scenarios, where spatial behavior under normal conditions may not typically apply (Wang and Taylor, 2014). With the advent and popularization of geopositioning technologies incorporated in electronic devices and the massive expansion of the digital world, the individual leaves behind traces of information —data shadows— available for use in



research projects. These data, where the individual becomes a sensor—or the bearer of the sensor— become citizen-sensor data or human-sensor data (Goodchild, 2007). The differentiation between active and passive citizen-sensor data is determined by the awareness of the individual about the use of the generated data by the sensor (Birenboim and Shoval, 2016). Thus, active citizen-sensor data in spatial behavior studies is bound to the monitoring of participants with GNSS-loggers (Global Navigation Satellite System receivers) and other kind of spatial sensor such as accelerometers. Studies conducted exploiting active citizen-sensor data are common in disciplines such as epidemiology (Vazquez-Prokopec et al., 2009; Stanton et al., 2017), mobility and transportation (Wang et al., 2015) and tourism (Shoval et al., 2011), but their application in stressed situations (emergencies) remains largely unexplored.

In contrast, passive citizen-sensor data rely on the digital data shadow that individuals generate in their daily life. The most widely used passive citizen-sensor data sources in spatial behavior applications are mobile telephone, smart card, Wi-Fi and Bluetooth and social media. The data needs for the study of spatial behavior in different emergencies/hazards is event dependent, which complicates providing an overall picture of the strengths and pitfalls of each data source. For instance, what could be considered low spatial accuracy for a terrorist attack event might suffice for a large-scale evacuation such as a hurricane event.

Mobile telephone data is a rich source of spatial information that provides representative samples (as cell-phone penetration is virtually 100%), and good spatial and temporal resolution (González et al., 2008; Birenboim and Shoval, 2016), which have attracted the attention of numerous fields such as sociology (Amini et al., 2014), public health

(Wesolowski et al., 2012), transportation (Çolak et al., 2016), geography (Phithakkitnukoon et al., 2012) and urban studies (Bajardi et al., 2015). The use of mobile phone data is limited in hazards. Only epidemiologists focusing on biohazards have systematically collected and analyzed mobile phone data to infer human mobility patterns (Wesolowski et al., 2012; Vogel et al., 2015). Others have leveraged phone call data to understand the after-event migrations of a disaster caused by an earthquake (Bengtsson et al., 2011) or more recently by a hurricane (Echenique and Melgar, 2018). Data accessibility (collected and owned by private corporations) and the lack of information about the individuals (normally bounded to privacy agreements) are the major reported limitations of these data (Widhalm et al., 2015).

Smart card data refers to any pocket-sized card with embedded integrated circuits, which is a common feature nowadays in finances, communications, personal identifications, public transportations, health care etc. Mobility studies leveraging the georeferenced data shadow from smart cards explore, for instance, the spatial behavior of foreign tourists (Sobolevsky et al., 2014) or the spatiotemporal dynamism of the urban activity spaces (Gong et al., 2017). Smart card data offers a refined spatial resolution — although bound to the receptor's location— with a variable temporal resolution depending upon the type of data pursued (Zhong et al., 2015; Gong et al., 2017). The representativeness of a sample derived from smart card data is dependent upon the specific data analyzed, but the dissemination of these cards across today's society has the potential of granting representative enough samples. Data accessibility and semantically poor information is again the main constraint of this approach although, similar to phone data, researchers have been able to sign data usage agreements with private corporations.

Wireless networks such as Wi-Fi and Bluetooth have a potential application for spatial behavior research (Perttunen et al., 2014). A singular limitation of these data is its reduced spatial coverage, as the geolocation of the individuals is only possible within the range of the devices' signal (Kontokosta and Johnson, 2017). To this, we must add the difficulty of accessing the data, the lack of information of the device bearer (which complicates the understanding of how representative a sample derived from these data is), and privacy and confidentiality concerns (Kontokosta and Johnson, 2017). On the positive side, the possibility of real-time positioning and the capacity of locating devices indoors make these technologies potentially interesting for applications in hazards such as body search after building collapses (earthquakes, avalanches, landslides etc.) or terrorist attacks (Cooper et al., 2016; Moon et al., 2016).

Finally, the advent of social media data has meant a revolution for research, with applications in a number of disciplines (e.g. Widener and Li, 2014; Gerber, 2014; Zhai et al., 2015; Li et al., 2018b). Easily accessible data has largely benefitted emergency management and crisis communication (e.g. Ukkusuri et al., 2014; Spence et al., 2016). Twitter has become the largest source of social media data for research due to its policy of free access to 1% of its total content (Burton et al., 2012), whereas other platforms such as Facebook remain virtually unexploited because of the stricter privacy policy about data sharing. For instance, Wang and Taylor (2014, 2016) disclosed spatiotemporal patterns of human mobility in different disasters by leveraging geotagged tweets. Dong et al. (2013) proposed a method to automatically visualize collected live disaster data from Twitter using word clouds, spatial maps and dynamic online activity graphs. Martín et al. (2017) advanced an approach for leveraging Twitter to estimate evacuation compliance, then

applied to other events such as Hurricane Sandy (Kumar and Ukkusuri, 2018). These studies revealed the strengths of these data: ubiquitous within the digital world, increasingly predominant in today's society and particularly within the younger age segments of the population; close to real-time data; and simple and free data accessibility, contingent upon enough computing power and programming skills. The weaknesses of social media include its low spatial resolution —as only a small fraction of the tweets is georeferenced with exact coordinates, the restricted periodicity or temporal resolution, which led Andrienko et al. (2012) to define it as sparse and episodic, and potential biases (Martín et al., 2017).

### 3.3. DISASTER CONTEXT AND STUDY AREAS

The study region comprises two coastal regions recently affected by Hurricane Matthew (2016) and Hurricane Irma (2017). This includes the coastal counties of South Carolina during Matthew —Beaufort, Berkeley, Charleston, Colleton, Dorchester, Georgetown, Horry and Jasper, and the most affected counties of Florida during Irma —Broward, Charlotte, Collier, Hillsborough, Lee, Manatee, Martin, Miami-Dade, Monroe, Palm Beach, Pasco, Pinellas and Sarasota— (Figure 3.1).

Hurricane Matthew was the strongest storm recorded in the Atlantic during the 2016 season, the first Category 5 hurricane since 2007 in this basin (Stewart, 2017). Matthew formed on September 28, 2016, and rapidly intensified, reaching the category 5 status on the Saffir-Simpson scale three days after. The storm became extratropical near the coast of North Carolina on October 9. It caused a total of 585 direct deaths: 546 in Haiti, 34 in the United States (Florida – 2, Georgia – 2, South Carolina – 4, North Carolina – 25, and Virginia – 1), 4 in the Dominican Republic, and 1 in St. Vincent and the Grenadines

(Stewart, 2017). Given the track and intensity of the storm, coastal residents from Florida to North Carolina were advised to evacuate in advance of the storm.

Hurricane Irma was a category 5 hurricane that caused over 130 deaths (44 in the Caribbean and 90 in the United States) and whose damage estimates places it as one of the costliest hurricanes on record with an estimate of 50 billion USD (Cangialosi et al., 2018). Formed on August 30, 2017, it battered Barbuda, Saint Barthélemy, Saint Martin, Anguilla, and the Virgin Islands as a Category 5 hurricane. Following a westward track, it later affected Cuba before turning north and making landfall in the Florida Keys as a category 4 (September 10). It finally lost tropical force winds over Georgia on September 12. The threat posed by Irma prompted one of the largest evacuations in the recent U.S. history with millions of people along the coasts of Florida, Georgia and South Carolina mandated or advised to seek refuge away from home (Ellis and Levenson, 2017; Shepherd, 2017; Wilks, 2017).

### 3.4. METHODS

#### 3.4.1 DATA COLLECTION AND PREPROCESSING

##### 3.4.1.1 SURVEY DATA

The Hazards and Vulnerability Research Institute (HVRI) conducted a post-hurricane survey mailing 19,829 postcards to residents along coastal counties of South Carolina to understand the response to mandatory evacuation orders issued in advance of Hurricane Matthew (Table 3.1). Infogroup®, a data analytics services firm provided the random addresses. The postcards included a link to an online survey platform by SurveyMonkey®. The survey consisted of standard behavioral hurricane evacuation questions covering general demographic information, prior experience, risk area awareness, information

sources and influences on behavior. The intent of the method used to deliver the survey (i.e. online survey) was to accelerate the data entry process and to minimize the costs. The survey was accessible from November to mid-December 2016 and obtained 329 responses (less than 2% response rate), of which 313 responses were valid for the purposes of this study. Several factors might have contributed to this low response rate. For instance, the web-based format we used to conduct the survey traditionally has a low response rate (Fan and Yan, 2010). In addition, the coincidence in time of the survey with the presidential election of 2016 and holiday season might also relate with the low response. HVRI did not conduct any survey for Hurricane Irma evacuations.

#### 3.4.1.2 TWITTER DATA

Over 1.4 billion geotagged tweets from the entire world comprising two six-month periods from July 1st, 2016 until December 31st, 2016 and from July 1st, 2017 to December 31st, 2017 were collected using the Twitter Stream Application Programming Interface (API). These tweets are stored and managed in a Hadoop (<http://hadoop.apache.org>) environment that served as tweet repository for this study. The repository was then queried with Apache Impala (<https://impala.incubator.apache.org>) using spatiotemporal and user information criteria.

Following Martín et al. (2017), we collected one pre-evacuation dataset between October 2nd and October 4th, 2016 for Hurricane Matthew and another for Hurricane Irma between September 3rd and September 5th, 2017. To obtain the largest possible number of users in the counties of study prior to the evacuations the keyword of the queries was left blank (Table 3.2). The minimal spatial accuracy level needed was city-level since the pursued aggregation was county level. The limitation of a county-level aggregation derives

from the fact that only around 15% of geo-located tweets have spatial accuracy finer than the city level, precluding from obtaining representative samples in more detailed scales. Looking at the tweet source —Twitter provides this information for each tweet— we were able to filter out non-human (bot) users. For instance, we did not include tweets from sources such as Tweetbot for IS, Tweetbot for Mac, and TweetMyJOBS. The non-human character of these users was manually confirmed analyzing the nature of their tweets (automated weather reports, job offers, advertising etc.). Later, the tweet repository was queried with an open-ended region (world) with the active twitter users identified in the pre-evacuation stage—those who tweeted at least one time. The timeframes for the post-evacuation periods assumed that by October 7th, 2016 at 6 pm (Hurricane Matthew in South Carolina) and September 5th, 2017 (Hurricane Irma in Florida) the evacuations concluded, taking into consideration the night (Dow and Cutter, 2002) and rainy conditions (first effects of Matthew and Irma, respectively). In a similar manner, we assumed the return period had not been initiated by October 9th, 2016 at 10am (Hurricane Matthew in South Carolina) and September 11th, 2017 at 8am (Hurricane Irma in Florida) considering that the trip would have implied driving at night and in rainy and flood conditions. Thus, if a user active in the coastal counties during the pre-evacuation period tweeted from outside these counties during the post-evacuation, we considered the user as an evacuee (Table 3.3).

Once the active users in both the pre-evacuation and the post-evacuation periods were identified, we collected their background information (residential status, gender, age, race/ethnicity). We determined the residential status based on the median center of each user's tweets during a 6-month period from July to December (in the year of the respective

hurricane), excluding two major holiday seasons (Thanksgiving and Christmas) where vacation movements might introduce noisy information and hinder home location identification. If the user's main tweeting location (based on their geotagged tweets) fell within the studied counties for each hurricane, we categorized them as a resident. Conversely, if the median of the location of tweets was outside of these counties, the user was a non-resident. We controlled the validity of this method by comparison with the user supplied home location. The literature is rich in methods for predicting users' demographic features at large scale —matching usernames to public data sources, supervised learning approaches exploiting tweet content etc. (Culotta et al., 2016). In our case, the size of the sample is small enough to assure a high-confidence demographic characterization by visiting the user's public profile. Profile pictures, usernames, full name, description, URLs, multimedia content and tweets uploaded by the users served to manually estimate gender (female or male), an age range (17 years or younger, 17-24 years, 25-34 years, 35-44 years, 45-54 years, 55-64 years, 65-74 years, and 75 years or older), and a race and/or ethnic group (white, black, hispanic or latino, and others). Table 3.4 further details these race/ethnic categories.

### 3.4.2. STATISTICAL ANALYSIS

In order to test the association between the sociodemographic characteristics of the sampled population and their evacuation behavior we conducted bivariate Chi-Squared tests of independence. This test determines the relationships between two categorical variables. The dependent variable is a binary signaling the evacuation behavior of the individuals (evacuated or non-evacuated) (Table 3.5). In the study we tested the association



of five independent variables including data method (2 groups), gender (2 groups), age (7 groups), race (4 groups), and residential status (2 groups) (Table 3.5).

### 3.5. RESULTS AND DISCUSSION

#### 3.5.1. INTEGRATION OF SURVEY DATA AND TWITTER

There is always a concern about the representativeness of survey questionnaires based on identified biases such a whiter and older sample in comparison to the overall population. Johnson and Wislar (2012), for example, warned of low sample sizes due to of decreasing response rates. This was observed in our survey, which obtained a response rate of 1.7% (329 respondents), confirming that online questionnaire surveys might be an ineffective method to study evacuation behavior, despite its low cost to administer. On the other hand, the Twitter approach obtained 1,760 users for Hurricane Matthew.

Table 3.6 presents the demographic profile of the web-based survey and the Twitter sample and allows a comparison with the demographic profile of the coastal counties of South Carolina. Beginning with the gender distribution of the samples, we observed fairly good matches with the overall population distribution, only slightly underrepresenting females. However, Twitter does show an asymmetry in the gender distribution. The sample underrepresents males (40.3%), which diverges from previous demographics of Twitter users suggesting an over-representation, or a male bias (Mislove et al., 2011) but confirms recent findings by Jiang et al. (2018). The race/ethnicity distribution of the survey reveals a racial bias towards the white population (86.3%), thus underrepresenting minorities such as blacks or hispanic/latino populations, a finding equally observed in Jiang et al. (2018). While the Twitter sample data are closer to the overall population distribution (66.7% white

population vs. 72.1% white Twitter population), a bias still exists towards the racial majority (whites).

Figure 3.2 and Figure 3.4 visually display the age and gender structure of the samples compared to the overall population using population pyramids. The base of the pyramid is well represented by the Twitter sample, bearing in mind that the youngest Twitter users are middle school and high school students and the population percentages include individuals from 0 to 17 years old. A female bias is in this age range and continues to the 18-24 cohort before reversing in the 35 to 54 age groups, which agrees with Leak and Lansley (2018). The 18-24 and 25-34 age cohorts are overrepresented in Twitter, showing an obvious bias towards younger populations. For surveys, the representation is almost negligible under 44 years, with a very low number of participants. The survey method gains relevance as Twitter decreases in older cohorts, reaching its maximum in the 65-74 year range. Together, the combination of survey and Twitter has three gaps in the age/gender distribution compared to the overall population in the study area: the 45-54 female cohort; 55-64 males and females; and females 75 years and over.

### 3.5.2. HURRICANE MATTHEW IN SOUTH CAROLINA

The Chi-square tests indicate that there was a significant association between residential status ( $\chi^2 (1, N=2073) = 195.996, p<0.001$ ) and age ( $\chi^2 (6, N=1650) = 33.647, p<0.001$ ) with the evacuation behavior of residents in the coastal counties of South Carolina during Hurricane Matthew (Table 3.7). This association is strong for residential status ( $\phi = -0.307$ ), however, when controlling the effect of the sample size, the level of association between age and evacuation behavior is weak (Cramer's  $V = 0.143$ ). Previous literature on the effect of age on evacuation behavior is inconclusive. Baker (1991) in his review of

evacuation studies found that most failed to establish an association between age and behavior while Huang et al. (2016a) reported that 59% of the studies they reviewed had non-significant correlations as well. Our results differ from these meta-analyses and agree with those who support the importance of age in hurricane evacuation behavior (Van Willigen et al., 2005).

Using residential status as another independent variable, we did not find any comparable study that assessed the evacuation behavior of residents and non-residents at the same time. The difficulty in reaching non-residents in the aftermath of an evacuation is likely behind the paucity of studies on this behavioral correlate. Twitter offers a possibility to overcome this; however, we advise caution in analyzing the results as we cannot infer the motivation of people's movements. For instance, a non-resident could be listed as evacuated because of leaving the coast due to the end of a vacation/work period, even though there was no real intentional protective action from a threat.

We found no association between the method —Twitter or survey— and evacuation decision-making ( $\chi^2(1, N=2073) = 2.352, p=0.125$ ), which is confirmed adjusting by the sample size ( $\phi = -0.034$ ). We believe this is a positive result confirming the validity of Twitter for studying selected factors related to evacuation behavior. Gender was also not associated with evacuation behavior ( $\chi^2(1, N=1962) = 0.890, p=0.345$ ). There is a general lack of agreement in the relation of gender and evacuation decision-making. Huang et al. (2016a) review reveals that 44% of the evacuation studies reported non-significant correlations between gender and evacuation behavior, however other analyses have reported weak associations (Huang et al., 2016b). Marginal evidence was found between

race/ethnicity and the response to the evacuation order ( $\chi^2(3, N=1932) = 7.784, p=0.051$ ), however this association is trivial in magnitude (Cramer's  $V = 0.063$ ).

Taking a more granular look into the behavior of the different groups identified we observe how Twitter users were less likely to evacuate than survey respondents by a factor of 0.8, although this result is not significant ( $p > 0.05$ ) (Table 3.8). Residential status significantly affects the likelihood of evacuation and non-residents were 18.9 times more likely to evacuate than residents ( $p < 0.001$ ). Looking only at the residents, we find differences in the evacuation response of residents from different counties. Beaufort residents were 3.8 times more likely to evacuate than residents from the remaining 7 counties ( $p < 0.001$ ). This pattern coincides with the most recent hurricane evacuation study in South Carolina (Cutter et al., 2011), that concluded that the largest participation in an evacuation would take place in Beaufort County. The track and intensity of Matthew, approaching the state from the south, is likely to be partly responsible of this fact, especially if we consider that residents from other southernmost counties (Jasper and Colleton) were 1.7 and 1.4 more likely to evacuate respectively than the remaining counties, although these results are not statistically significant ( $p > 0.05$ ). Dorchester has the lowest storm surge potential risk of the 8 counties in the study area, which is probably the reason that explains that its residents were less likely to evacuate than their coastal-counties counterparts by a factor of 0.5 ( $p < 0.01$ ). Following this trend, evacuation rates decreased as we move northwardly along the coast, concurring with Cutter et al. (2011) findings. Thus, Georgetown and Horry's evacuation rates during Hurricane Matthew were lower in comparison with the rest of the counties.

Women are more likely to evacuate than men (Lindell et al., 2005; Dash and Gladwin, 2007). Bateman and Edwards (2002) examined the causes of this evidence and found it linked to socially constructed gender differences in care-giving roles, access to evacuation incentives, exposure to risk, and perceived risk. Our findings observe slightly higher likelihoods (odd ratio = 1.1) of evacuation of females, although the result is not significant ( $p>0.05$ ). While race/ethnicity as an overall predictor is not significant (Table 3.8) we found disparities in the behavior of different groups (Table 3.9). The white population were less likely to evacuate than the rest of the racial/ethnic groups by a factor of 0.8 ( $p<0.05$ ). We also found the hispanic/latino population were 1.9 more likely to evacuate than the remaining groups ( $p<0.05$ ). This higher likelihood of evacuation among hispanics/latinos agrees with studies such as Dixon et al. (2017). Others and blacks recorded non-significant ( $p>0.05$ ) higher evacuation rates than whites, which coincides with Huang et al. (2016a) conclusions in the behavior of these groups.

Figure 3.3 demonstrates that where Twitter and survey converge and sample sizes are sufficient, evacuation rates by age group tend to follow the same pattern. The 18-24 years cohort was 1.8 times more likely to evacuate ( $p<0.001$ ) than the remaining groups. This finding coincides with previous literature. For instance, Van Willigen et al. (2005), reported that evacuation rates within the 18-25 years group prior to the arrival of Hurricane Floyd in 1999 were significantly higher than older populations. This age group is mainly composed of students with a certain degree of independence. We believe that universities' closures —and therefore time availability and lack of personal obligations, fewer and weaker ties to the community, a shorter evacuation experience, and the recent memory of

the catastrophic South Carolina floods of 2015 led to a massive compliance with the evacuation order of this age group.

Sixty-six percent of young professionals (age group 25-34 years) complied with the evacuation order and were 1.2 times more likely to evacuate than the remaining age groups, although this result is not statistically significant ( $p>0.05$ ). The analysis confirmed middle-aged adults (35 to 64) as the group more prone to stay and ride out the storm within their communities. However, the association in the 55-64 interval is not significant ( $p>0.05$ ) as the sample size is limited. After retirement age, evacuation likelihood increased progressively (not significant,  $p>0.05$ ), which agrees with findings in Bowser (2013), who analyzed the evacuation behavior of the elderly in South Carolina. Nevertheless, our findings contradict most studies on the evacuation behavior of the elderly, which suggested that this group were less likely to evacuate (Gladwin and Peacock, 1997) despite their vulnerability during disasters (Hutton, 2008).

### 3.5.3. HURRICANE IRMA IN FLORIDA

Florida's evacuations differ from evacuations in other U.S. states due to its morphology and the relative high frequency of tropical cyclones (Elsner et al., 2004). Thus, officials tend to issue evacuations at a finer scale—if we compare with South Carolina—and often only those living in low-lying areas close to the shoreline are mandated to leave. In addition, as a response to Hurricane Andrew's extended damage in 1992, Florida dictated a more restrictive building code thought to make structures endure hurricane-force winds (Gurley et al., 2006). Consequently, many Floridians decide to hunker down at home while some of the ones that choose to evacuate stay with family or friends within their own county. Twitter's finest spatial resolution (providing a large enough sample size) is city-

level, therefore obviating all these intra-county evacuations. As earlier discussed, the use of other sources of passive citizen-sensor might improve our understanding of evacuations at a finer scale and therefore produce more accurate results.

Table 3.9 presents the statistical significance of the association between evacuation behavior and the four independent variables. Note we did not conduct a survey for Hurricane Irma in Florida and therefore results come from Twitter alone. Out of the four variables, only gender was not statistically associated to evacuation behavior ( $\chi^2 (1, N=8936) = 0.431, p=0.512$ ). In other words, males and females evacuated at roughly the same rate (Table 3.10)

Similar to Matthew case study, residential status was strongly associated with evacuation behavior ( $\chi^2 (1, N=10017) = 2427.962, p<0.001$ ) during Hurricane Irma in Florida, even when accounting for the effect size ( $\phi = -0.492$ ). This finding confirms that despite non-residents increased vulnerability driven by the potential lack of knowledge about local hazards and material resources or social networks to undertake protective actions (Phillips and Morrow, 2007; Villegas et al., 2013), temporary residents do find their way out of harm's way and were 22.2 times more likely to evacuate than residents, which is consistent with Matyas et al. (2011) and Cahyanto et al. (2014). Looking at the breakdown of the residents' evacuation response by county (Table 3.10), we find that those affected by Irma's most intense winds —Monroe, Collier, Lee, Charlotte— recorded the highest evacuation rates. For instance, residents of Monroe County, which includes the Florida Keys (severely affected by Irma), were 4.5 times more likely to evacuate than residents from the rest of the study area. Conversely, on the Atlantic coast, Broward County had an evacuation rate of 17.3%. An evacuation rate of local residents of 23.2% might

seem low considering the threat of a category 4 hurricane, however, given the nature of mandatory evacuations in removing residents from storm surge inundation and flooding, not the effects of winds, this makes some intuitive sense.

In contrast with Matthew in South Carolina and Collins et al. (2018) study on Irma evacuation in Florida, we saw a significant ( $\chi^2 (5, N=8389) = 102.651, p<0.001$ ) weak (Cramer's  $V = 0.111$ ) association between the race/ethnicity of the individuals and their predisposition to evacuate. The discordance with Collins et al. (2018) might be due to the different study areas, as their study only focused on Pinellas and Hillsborough counties while this research includes several additional counties. Regarding the difference in the association of race/ethnicity with evacuation behavior between South Carolina and Florida, we believe that it is possible that there exists place-specificity (spatial heterogeneity) in the factors affecting population's evacuation decision-making, an idea also suggested in Lazo et al. (2015). In other words, what holds true for one area might not be a factor in another. Thus, during Hurricane Irma, Hispanic/Latino were significantly less likely to evacuate (by a factor of 0.7,  $p<0.001$ ) and only 27.5% chose to leave their county. In this case, white residents were 1.3 times more likely to leave the county before the arrival of the hurricane ( $p<0.001$ ). The others group also had a positive association with evacuation behavior and were 2.4 times more likely to evacuate ( $p<0.001$ ).

Lastly, age was again a significant predictor variable of evacuation behavior ( $\chi^2 (6, N=7513) = 41.566, p<0.001$ ), however, when accounting for the effect size, this association is very weak. As seen with Matthew, we find a sample biased towards the young and with limited sample sizes over 55 years. We observe the same pattern of a highly female-skewed sample under 25 years that reverses progressively (Leak and Lansley, 2018). The largest



evacuation rate was recorded among individuals from 25 to 34 years (38.2%) and gradually decreased with age, being particularly low over 65 years old. This pattern agrees with literature suggesting the elderly lower predisposition for evacuation (Gladwin and Peacock, 1997) but stresses the differences found between the behavior of different subpopulations in South Carolina and Florida.

### 3.6. CONCLUSIONS AND FURTHER RESEARCH

This article responds to calls for more creative methods to understand current spatial behavior related to hurricane evacuations (Baker, 2009). It bridges traditional methods of data collection for evacuation behavior (surveys) with the opportunity created by Big Data for studying massive amounts of information to infer the behavior of large populations. The study contributes to the literature in two ways. First, it focuses on developing the method advanced by Martín et al. (2017) by applying it to a different event and comparing it to a traditional approach. This allows identifying the opportunities and challenges that the exploitation of passive citizen sensor data brings to evacuation behavior assessments. Secondly, it advances the knowledge of the field by focusing attention in under-studied groups such as the young, racial minorities, and the non-residents.

One of our main findings is that geotagged tweets can complement questionnaire surveys and create a more representative sample of the overall population in terms of age and race/ethnicity distribution. However, we still found gaps where the sample needs improvement. Particularly, the age cohort from 55 to 65 years is poorly represented by the combination of survey and Twitter. It is likely that the representativeness of older cohorts will improve as tech-savvy generations get older and Twitter users are more evenly present throughout all segments of the society. Nevertheless, we believe incorporating other

sources of passive citizen-sensor data could improve the challenges and biases we encountered, such as the inability of Twitter to discriminate intra-county evacuations. Whether Twitter users can fairly represent a population is still debatable and requires further analyses that determine and quantify potential biases (Martín et al., 2017), however some authors have highlighted that existing experimental research in the social sciences has often utilized samples that are more flawed than social media samples (Spence et al., 2016).

In this study, we focused on three independent variables widely used to predict evacuation behavior (gender, age, and race/ethnicity) and we explored a fourth factor, residential status. In general, we can confirm that gender played no role in these two hurricane evacuations whatsoever, and that residential status were both significantly relevant. Results on age and race/ethnicity indicate differences in the behavior of the residents in the coastal counties of South Carolina during Hurricane Matthew and the studied counties of Florida during Hurricane Irma, which might indicate place-based differences in the predictors. In subsequent studies, Twitter might also be utilized — exploiting self-provided information and/or content-based mining— to infer and test other variables such as marital status, presence of children in the household, or estimations of education attainment and income ranges. The method would also permit tracking the same Twitter users in different events and thus recreating their behavior based on storm track and intensity or official warnings. The analysis here presented and future advancements in this line of research might be useful for authorities to both better prepare evacuation campaigns — trying to reach subpopulations with particularly low evacuation compliance rates, and if this approach can be automated to offer close to real time information, the

dynamic hurricane evacuation response might assist managers to improve their response and minimize the obstacles of both the evacuation and return processes.

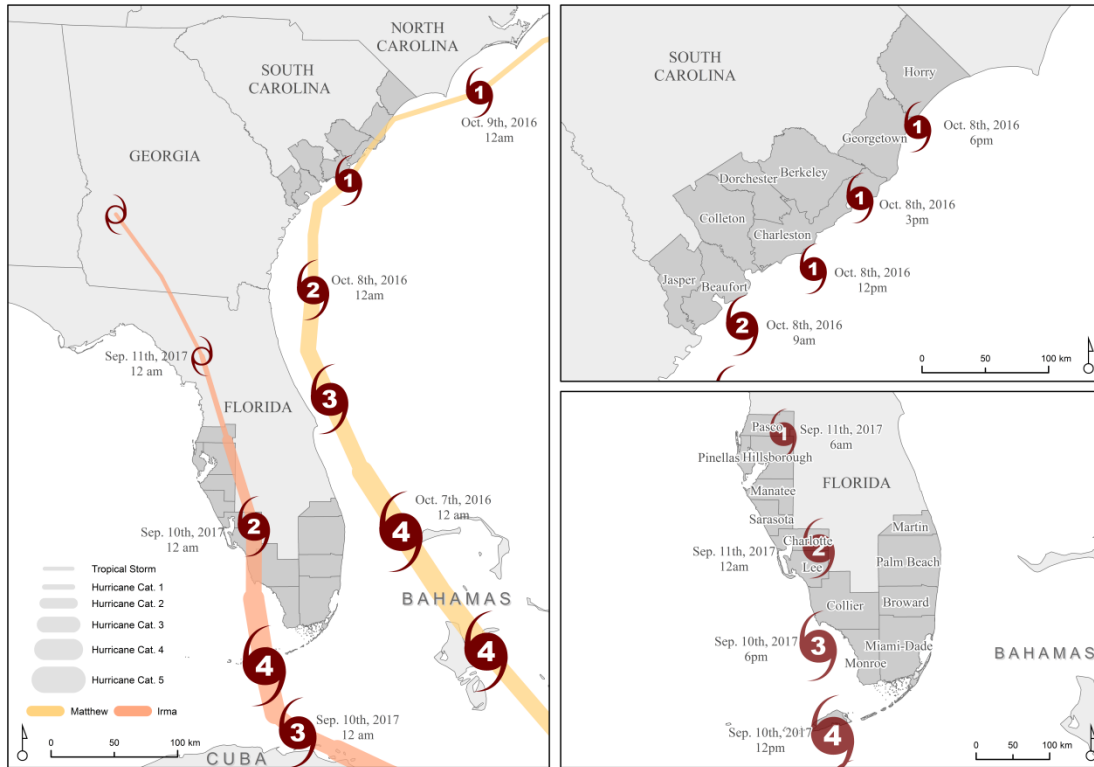


Figure 3.1. Area of study and storms tracks.

Table 3.1. Survey response rates and confidence interval

Postcards delivered	19,829
Responses received	329
Response rate (%)	1.66
Responses valid	313
Total population in study area	973,340
Confidence interval of all responses received	$\pm 5.40\%$

Table 3.2. Dataset query conditions

<b>Dataset name</b>	<b>Region</b>	<b>Keywords / Selection criteria</b>	<b>Time period</b>	<b>Location accuracy</b>
Matthew Pre-evacuation	Beaufort, Berkeley, Charleston, Colleton, Dorchester, Georgetown, Horry and Jasper	*	10/02/2016 - 10/04/2016	City, Neighborhood, Point
Matthew Post-evacuation	World	Active users in Matthew Pre-evacuation	10/07/2016, 6pm - 10/09/2016, 10am	City, Neighborhood, Point
Irma Pre-Evacuation	Broward, Charlotte, Collier, Hillsborough, Lee, Manatee, Martin, Miami-Dade, Monroe, Palm Beach, Pasco, Pinellas and Sarasota	*	09/03/2017 - 09/05/2017	City, Neighborhood, Point
Irma Post-Evacuation	World	Active users in Irma Pre-evacuation	09/09/2017, 8pm – 09/11/2017, 8am	City, Neighborhood, Point

Table 3.3. Twitter active users in the study areas

	<b>Active Twitter Users</b>	<b>Evacuated</b>	<b>Non-Evacuated</b>
Hurricane Matthew - SC	1,760	1,101	659
Hurricane Irma - FL	10,017	3,324	6,693

Table 3.4. Race/ethnic categories

<b>Race/Ethnicity</b>	<b>Description</b>
White	A person whose facial features resemble any of the original peoples of Europe
Black	A person whose facial features resemble any of the original peoples of the black racial groups of Africa
Hispanic or Latino	A person whose country of origin is in Central America, the Caribbean or South America and whose native language is Spanish or Portuguese
Others	A person whose facial features resemble any of the original peoples of the Far East, Southeast Asia, India, Middle East or North Africa

Table 3.5. Dependent and independent variables for evacuation behavior analysis

<b>Dependent variable</b>		<b>Independent variables</b>	
<b>Variable</b>	<b>Categories</b>	<b>Variable</b>	<b>Categories-Intervals</b>
Evacuation	Evacuated/Not evacuated	Method	Survey / Twitter
		Gender	Male/Female
		Age	18-24/25-34/35-44/45-54/55-64/65-74/>75
		Race	White/Black/Hispanic-Latino/Others
		Residential status	Resident / Non-resident

Table 3.6. Demographic profile of survey and twitter samples (Hurricane Matthew).

	Population coastal counties SC <sup>1</sup>		Survey		Twitter	
	Total	Percentage	Total	Percentage	Total	Percentage
<b>Gender</b>						
Female	680,896	51.2%	148	47.3%	952	54.1%
Male	648,983	48.8%	152	48.6%	710	40.3%
Unknown			13	4.2%	98	5.6%
<b>Race/Ethnicity</b>						
Black	309,655	23.3%	16	5.1%	286	16.3%
Hispanic/Latino <sup>2</sup>	79,250	6.0%	2	0.6%	53	3.0%
White	886,555	66.7%	270	86.3%	1269	72.1%
Other	25,769	1.9%	12	3.8%	24	1.4%
Unknown <sup>3</sup>	28650	2.2%	13	4.2%	128	7.3%
<b>Age</b>						
17 years or younger	282,760	21.3%	0	0.0%	268	15.2%
18-24 years	121,699	9.2%	1	0.3%	360	20.5%
25-34 years	186,959	14.1%	12	3.8%	605	34.4%
35-44 years	161,370	12.1%	15	4.8%	207	11.8%
45-54 years	172,501	13.0%	32	10.2%	122	6.9%
55-64 years	177,679	13.4%	48	15.3%	39	2.2%
65-74 years	143,590	10.8%	146	46.6%	16	0.9%
75 years or older	83,321	6.3%	44	14.1%	3	0.2%
Unknown	0	0.0%	15	4.8%	140	8.0%

<sup>1</sup>. Source: 2012-2016 American Community Survey (ACS) 5-year estimates. <sup>2</sup>. Includes all reported races with hispanic/latino ethnicity. <sup>3</sup>. In ACS data, this includes two or more races. In the survey it includes not-reported race/ethnicity. In Twitter, it includes users whose race/ethnicity could not be assigned.

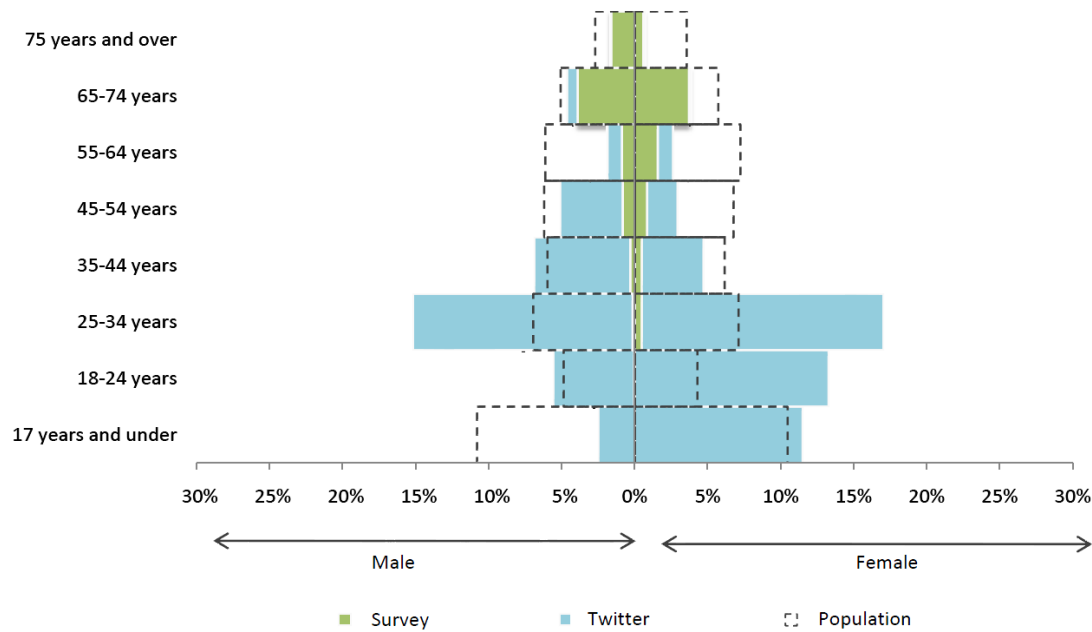


Figure 3.2. Comparison of age and gender distribution between survey and Twitter samples and the overall population of the coastal counties of South Carolina. Source: 2012-2016 ACS 5-year estimates.

Table 3.7. Statistical significance of predictors of evacuation behavior for Hurricane Matthew in the coastal counties of South Carolina

Predictor variable	N	$\chi^2$	p-value	Phi/Cramer's V	df
Method	2073	2.352	0.125	-0.034	1
Residential status	2073	195.996	< 0.001***	-0.307	1
Gender	1962	0.890	0.345	-0.021	1
Race	1932	7.784	0.051	0.063 <sup>#</sup>	3
Age	1650	33.647	<0 .001***	0.143 <sup>#</sup>	6

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

<sup>#</sup> Cramer's V



Table 3.8. Evacuation rates and odds ratio for Hurricane Matthew in the coastal counties of South Carolina

	N	Evacuated	Evacuation rate	Odds ratio	p-value
<b>Method (ref=Survey)</b>	2073				
Survey	313	210	67.10%		
Twitter	1760	1101	62.60%	0.8	0.125
<b>Residential status (ref=Resident)</b>	2073				
Resident	1717	969	56.40%		
(ref= remaining counties)					
Jasper	38	26	68.80%	1.7	0.132
Beaufort	208	169	81.30%	3.8	< 0.001***
Colleton	14	9	64.30%	1.4	0.552
Dorchester	107	44	41.10%	0.5	0.001**
Charleston	394	218	55.30%	0.9	0.614
Berkeley	161	82	50.90%	0.8	0.139
Georgetown	80	38	47.50%	0.7	0.099
Horry	715	383	53.60%	0.8	0.043*
Non-resident	356	342	96.10%	18.9	< 0.001***
<b>Gender (ref=Male)</b>	1962				
Female	1100	708	64.40%	1.1	0.345
Male	862	537	62.30%		
<b>Race/Ethnicity (ref=remaining groups)</b>	1932				
Black	302	200	66.20%	1.1	0.304
Hispanic/Latino	55	42	76.40%	1.9	0.046*

	White	1539	960	62.40%	0.8	0.028*
	Other	36	27	75.00%	1.7	0.152
<b>Age (ref=remaining groups)</b>		1650				
	18-24 years	361	266	73.70%	1.8	< 0.001***
	25-34 years	617	409	66.30%	1.2	0.121
	35-44 years	222	125	56.30%	0.7	0.013*
	45-54 years	154	81	52.60%	0.6	0.003**
	55-64 years	87	52	59.80%	0.8	0.421
	65-74 years	162	112	69.10%	1.3	0.141
	75 years or older	47	34	72.30%	1.5	0.218

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

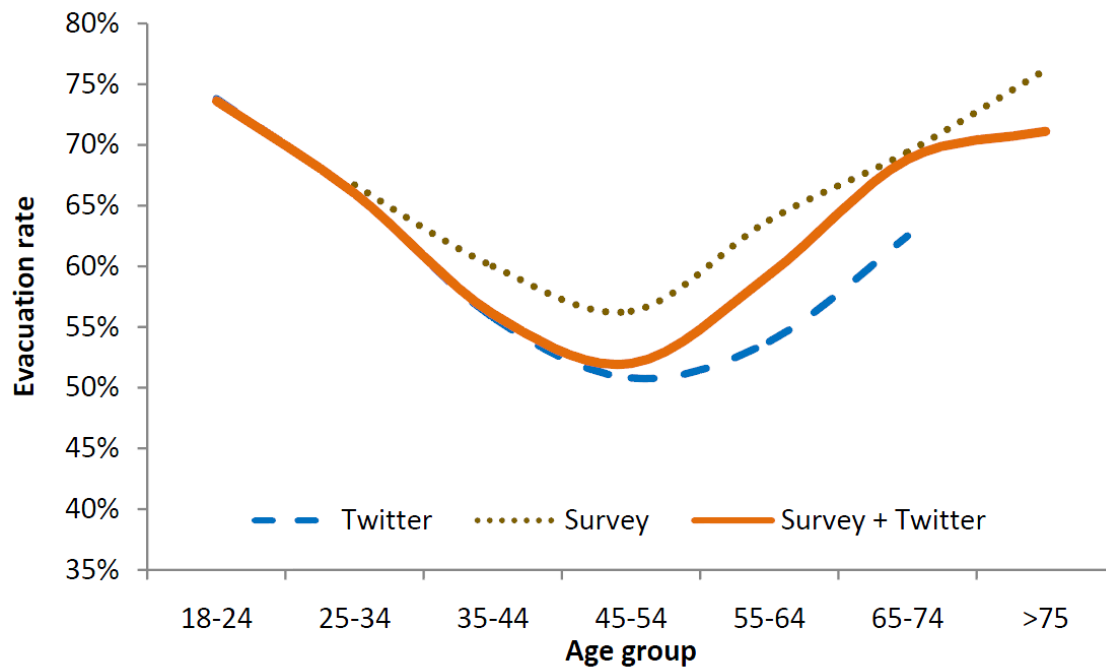


Figure 3.3. Evacuation behavior by age group and method. Survey results under 25 years were not included due to a small sample size (1 respondent). Twitter results over 75 were also excluded for this reason (3 users).

Table 3.9. Statistical significance of predictors of evacuation behavior for Hurricane Irma in Florida

Predictor variable	N	$\chi^2$	p-value	Phi/Cramer's V	df
Residential status	10017	2427.962	<0.001***	-0.492	1
Gender	8936	0.431	0.512	-0.007	1
Race	8389	102.651	< 0.001***	0.111 <sup>#</sup>	3
Age	7513	41.566	< 0.001***	0.074 <sup>#</sup>	6

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

<sup>#</sup> Cramer's V

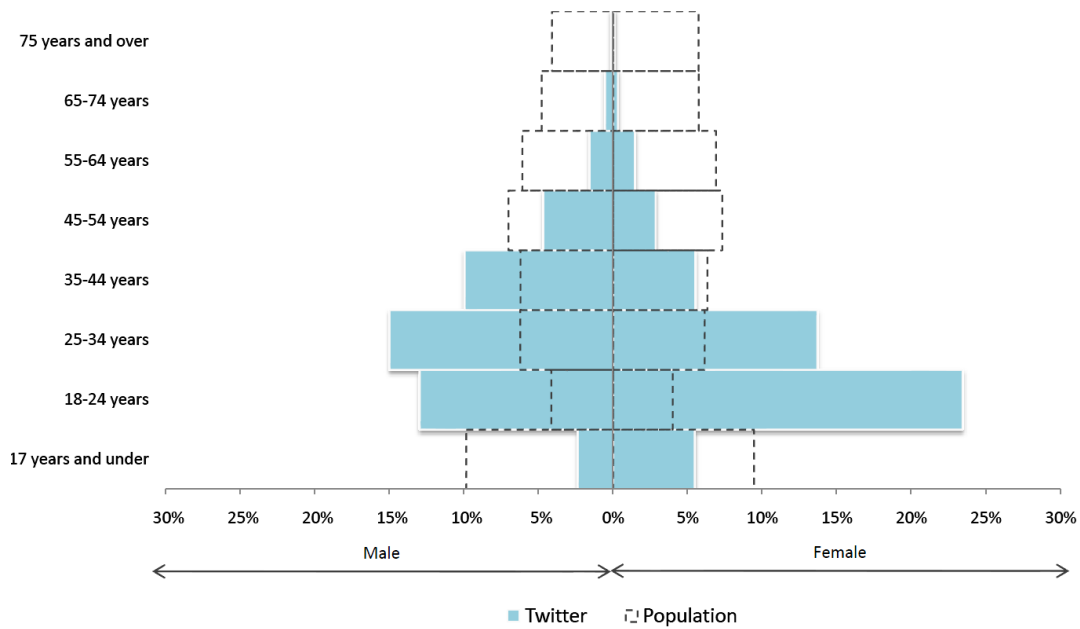


Figure 3.4. Comparison of age and gender distribution between Twitter sample and the overall population of the studied counties of Florida. Source: 2012-2016 ACS 5-year estimates.

Table 3.10. Evacuation rates and odds ratio for Hurricane Irma in Florida

	N	Evacuated	Evacuation rate	Odds ratio	p-value
<b>Residential status (ref=Resident)</b>	10017				
Resident (ref= remaining counties)	8451	1961	23.20%		
Martin	60	15	25.00%	1.1	0.741
Palm Beach	939	222	23.60%	1	0.736
Broward	1697	294	17.30%	0.6	< .001***
Miami-Dade	2653	600	22.60%	1	0.386
Monroe	38	22	57.90%	4.6	< .001***
Collier	49	21	42.90%	2.5	0.001**
Lee	312	112	35.90%	1.9	< .001***
Charlotte	31	10	32.30%	1.6	0.232
Sarasota	162	43	26.50%	1.2	0.309
Manatee	165	41	24.80%	1.1	0.613
Hillsborough	1276	363	28.40%	1.4	< .001***
Pinellas	721	156	21.60%	0.9	0.273
Pasco	348	62	17.80%	0.7	0.015*
Non-resident (ref=Internationals)	1566	1363	87.00%	22.2	< .001***
Rest of US	1236	1092	88.30%	1.7	0.003**
Internationals	330	271	82.10%		
<b>Gender (ref=Male)</b>	8936				
Female	4574	1541	33.70%	1	0.512
Male	4362	1441	33.00%		
<b>Race/Ethnicity (ref=remaining groups)</b>	8389				
Black	1656	551	33.30%	1	0.674

Hispanic/Latino	2091	670	27.50%	0.7	< .001***
White	4043	1471	36.40%	1.3	< .001***
Other	251	136	54.20%	2.4	< .001***
<b>Age (ref=remaining groups)</b>	7513				
18-24 years	2966	983	33.10%	1	0.43
25-34 years	2339	894	38.20%	1.3	< .001***
35-44 years	1260	461	36.60%	1.2	0.018*
45-54 years	613	182	29.70%	0.8	0.029*
55-64 years	246	62	25.20%	0.7	0.004**
65-74 years	70	15	21.40%	0.5	0.029*
75 years or older	19	4	21.10%	0.5	0.243

---

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

## **CHAPTER 4**

### **TRACKING THE DISRUPTION OF HURRICANE MARIA ON POPULATION MOVEMENTS IN PUERTO RICO THROUGH GEOTAGGED TWEETS<sup>3</sup>**

---

<sup>3</sup> Martín, Y., Cutter, S.L., Li, Z., Emrich, C.T., Mitchell, J.T. Tracking the disruption of Hurricane Maria on population movements in Puerto Rico through geotagged tweets. Submitted to *Population & Environment*.

#### 4.1 INTRODUCTION

Natural hazards have the power to change the environment of a region for years after an event. The destruction caused by a hurricane, an earthquake, or a tsunami can render some areas uninhabitable and force residents to relocate. Examples abound in recent times. For instance, the Boxing Day tsunami of 2004 in Southeast Asia displaced hundreds of thousands of residents, while in the United States, 2005's Hurricane Katrina caused many New Orleans and Mississippi residents to flee and never return. Increasing threats from sea level rise and climate change add uncertainty into the hazardous future of many coastal settlements (McGranahan et al., 2007; Curtis and Bergmans, 2018), not only increasing the number of people at risk, but more significantly those that may be displaced either temporarily or permanently.

Environmental displacement and migration are oft-studied concepts, however new research increasingly focuses on climate change (McLeman and Gemenne, 2018), and disasters as causal agents. The general study of human migration often encounters difficulties in form of the lack of accessible and reliable data (Willekens et al., 2016; Rango and Vespe, 2017), as documenting displaced or migrant populations often elude traditional methods, especially in developing countries, and even more so during post-disaster and emergency situations (Laczko, 2015). Despite reported improvements in the availability, quality, and comparability of migration data (Laczko, 2015), data concerns continue to constrain migration scholars (Spyratos et al., 2018). Some researchers have called for an increased collection of quantitative data to measure migration flows (Piguet, 2010; Bilsborrow and Henry, 2012) and for efforts to integrate more diverse, timely, and trustworthy information (United Nations, 2014). Big Data provide new possibilities to



tackle some of the limitations of traditional methods when tracking population movements. In particular, passive citizen sensor data such as geotagged social media hold an enormous potential for understanding spatial behavior in disaster situations. The literature demonstrates that Twitter is amenable for addressing some aspects of evacuation behavior (Martín et al., 2017; Kumar and Ukkusuri, 2018), but the study of post-disaster population movements is yet to be fully explored.

Responding to calls for innovative data-collection methods of population movements, we examine the suitability of Twitter data for the assessment of the disruption of population movements triggered by a disaster. Thus, we leverage Twitter data to explore the impact of Hurricane Maria on resident out-flows (displacement/migration) and non-residents inflows (tourism) in Puerto Rico. By individually analyzing the tweet location of 1,231 Twitter users, we estimate the total displacement, destinations, timing, and return of displaced Puerto Ricans. In addition, we test the association of gender, age, region of residence of the Twitter users with their displacement behavior. We contextualize our findings and the suitability of Twitter for assessing post-disaster displacement and migration through comparisons with recent studies about these processes in Puerto Rico such as Teralytics (2018), Sutter and Hernandez (2018), Hinojosa et al. (2018), Hinojosa and Meléndez (2018), and United States Census Bureau (2018a). For the analysis of population inflows into Puerto Rico, we compare the amount of non-resident Twitter users active every week during the post-disaster period (September 1, 2017 - August 31, 2018) to baseline pre-disaster levels (September 1, 2016 - August 31, 2017).

## 4.2 RESEARCH CONTEXT AND BACKGROUND

### 4.2.1. THE SCIENCE: POST-DISASTER POPULATION MOVEMENTS

Despite the lack of a thorough and sound conceptualization of disaster recovery (Johnson and Hayashi, 2012), disaster scholars recognize the multidimensional (demographic, infrastructural, economic, social, cultural, and psychological) nature of the post-event recovery process (Comerio, 2005; Chang, 2010). The population dimension of recovery is one essential element of the overall recovery picture, and it has increasingly attracted the interest of researchers, especially demographers, particularly after Hurricane Katrina and the devastation of New Orleans in 2005 (Fussell, 2015).

The definition and characterization of population movements is sensitive to spatiotemporal attributes. Population movements associated with either natural or anthropogenic hazards often start with the evacuation process —provided that the threat can be anticipated— and might extend several months or even years after the event, often involving post-disaster migration (Black et al., 2013). This temporal continuum creates confusion in the terminology employed by scholars approaching the field from diverse disciplines and backgrounds (e.g. sociology, geography, economics, or political science). Many have reported a lack of consistency in the use of concepts such as evacuees, displaced population, dislocated population, refugees, and migrants (Piguet et al., 2011; Mitchell et al., 2012). For instance, the term evacuee, often reserved for those who leave in advance of an incoming threat (Lindell et al., 2005), has also been used for those who decide or are forced to leave their homes in the aftermath of a disaster (Elliott and Pais, 2006). Variability across the temporal continuum also adds complexity, since an evacuee can transition to a

displaced person and a displaced person can become a migrant (Black et al., 2013; Adger et al., 2018).

In addition to departure time (pre-disaster versus post-disaster), other differences include the duration of the movement (ranging from short-term –temporary– to long-term or even permanent), volition (voluntary or involuntary movement), and spatial dimension (type of boundary crossed) (Fussell, 2012; McLeman, 2014). Regarding duration, there is no standard temporal definition of what constitutes a temporary resident (tourist, seasonal worker, displaced person) from a permanent migrant, and researchers and organizations have used several different thresholds —3 months, 6 months, 1 year— as identifiers (Bell et al., 2015). The distinction of the movement motivation is also fraught with ambiguity. Most research on this issue originates from studies about outmigration regions in developing countries (Oliver-Smith, 2009), where armed conflict (Melandar et al., 2009), hunger (Baro and Deubel, 2006) or disease (Toole, 1995) motivates massive population movements. Environmental-induced population movements are more complex to characterize. While some have focused on the effects of migrations on the environment (Bilsborrow, 2002), others put the focus on how —or if— the environment produces migrations (Hunter et al., 2015). Among the latter, several studies have attempted to link population movements to processes such as environmental degradation (Piguet, 2010) or to natural hazards such as drought (Gray and Mueller, 2012), tornadoes (Cross, 2014), or hurricanes (Fussell et al., 2017). For instance, a recent publication analyzing the impact of hurricanes on migrations into the United States from thirty Central American and Caribbean countries found a 6% increase in movements after the most damaging storms

(Spencer and Urquhart, 2018). However, others claim that this form of migration to the United States is understudied and in need of more attention (Mitchell et al., 2012).

The debate of hazard-induced population movements includes a well-established distinction between slow-onset environmental changes and rapid-onset environmental events. Slow-onset environmental changes such as drought or sea level rise typically result in progressive and long-term to permanent migrations (Laczko and Aghazarm, 2009; McLeman and Hunter, 2010), which some scholars claim is a coping mechanism that reflects adaptation to a changing environment (McLeman and Smit, 2006). Rapid-onset events such as hurricanes, earthquakes, or hazardous material spills often produce shorter-term and shorter-distance movements than slow-onset hazards, limiting international migration numbers to regular pre-disaster flows within the migration system (Laczko and Aghazarm, 2009; Findlay, 2011). Population movements responding to rapid-onset hazards are therefore often temporary, and displaced residents normally return to their communities when their community is restored (Curtis et al., 2015). Return migration is not the only population move in a recovery area (Fussell et al., 2014a). External investments and recovery funds sometimes revitalize the local economy and attract new residents to disaster-stricken places, a process described by Pais and Elliott (2008) as the “The Recovery Machine”. Post-disaster immigration spatially concentrates in the impact zones (re-construction) and in the urban development sector, altering the original demographic composition of the area with an influx of workers and new residents seeking opportunities (Pais and Elliott, 2008; Ehrenfeucht and Nelson, 2013).

Another variation in post-disaster population flows is the reduction in the number of tourists, an important economic mainstay for many tourism-based economies. Concerned

with this issue, studies by Faulkner (2001) and Ritchie (2008) encourage the adoption of proactive mitigation strategies to cope with the reduction of tourists in the aftermath of a disaster. Although most disasters induce an initial decrease in the arrival of national and international tourists, some events attract the curiosity of visitors and can indeed boost the tourism sector; a phenomenon labeled dark tourism or eco-disaster tourism (Gould and Lewis, 2007). Indeed, some disaster areas have been converted into tourist points of interest and have become a considerable source of revenue such as post-Katrina New Orleans neighborhoods (Pezzullo, 2009) and the National September 11 Memorial and Museum in New York City (Sather-Wagstaff, 2011).

No matter how it is discussed or how it affects recovery, one certainty remains, people evacuate during disasters. Some of them return and others do not, yet we as a society do not have a standard, replicable, and consistent way to tract a disaster diaspora, whether large or small.

#### 4.2.2. MEASURING POPULATION MOVEMENTS

##### 4.2.2.1 TRADITIONAL DATA SOURCES

The study of large-scale population flows commonly involves the use of population registers and/or censuses (Fussell et al., 2017) and surveys (Mallick and Vogt, 2014) as principal sources of migration data. Registers and censuses involve procedures for systematically collecting and recording information about a given population. Only a few countries regularly record vital statistics and residence changes for the whole population and collection rules are often-country specific (i.e. not compiling the same information), which makes registries less useful in many world regions. Nevertheless, where civil registries occur, researchers have an annual record of all migration events with great

geographic detail (Bell et al., 2015). Censuses are internationally accepted, and the United Nations recommends standards and methods to assist national statistical authorities in their compilation (United Nations, 2008). The strengths of censuses for the study of human migration embrace its full population and geographic coverage —universal sample and mandatory reply— and provide a rich demographic profile of the population.

On the downside, there are four broad concerns in using census information for migration studies. First, responsible public authorities often conduct censuses once every decade —as it requires considerable human and economic resources—and such a time lag between collection periods is insufficient for most migration research purposes (Fussell et al., 2014b). Second, censuses are not specifically designed to study aspects of migration and therefore can only include a limited number of questions about this issue. Third, researchers also have concerns about the reliability of the data (Bell et al., 2015). Fourth, undocumented or irregular migrants often are elusive to registries and censuses, which limits the usefulness of these approaches in the migration field (Laczko, 2015).

Surveys are typical sources of migration data and can present different designs: cross-sectional surveys, multiple cross-sectional surveys, and longitudinal or panel survey designs (Fussell et al., 2014b). Researchers can specifically model their questionnaires to investigate migration, which is the principal strength of the method. Thus, surveys are able to record detailed migration histories and motivations at comparatively lower expenses than censuses (Bell et al., 2015). Also, surveys are the preferred data collection method to relate environmental events to migration outcomes since the periodicity of censuses is often too sparse (Fussell et al., 2014b). Sampling decisions and resulting biases are the main

concern for researchers conducting surveys on migration, which relates to the inability to reach temporary migrants or migrants that already left the study area.

Tourism as a driver of population inflow is a multi-scale phenomenon studied from the tourist attraction itself to broader regional, national, and even global patterns. Depending upon the spatial scale of analysis, researchers use different methods to understand the behavior and patterns of tourism. Surveys and diaries are the main data-collection methods used in tourism research, especially at smaller spatial scales (Page and Hall, 2014). Accommodation and travel statistics also are conventional data sources for researchers (Page and Hall, 2014). However, cross-border and accommodation statistics often are too spatially coarse, temporally sparse, and semantically shallow to explore tourist decision-making (Raun et al., 2016). Also, some countries no longer collect some of these statistics (i.e. European Union member states). Surveying and self-report approaches are resource-demanding, difficult to apply in remote areas, and subject to biases (i.e. sample bias or recall bias) (Shoval and Ahas, 2016). Considering this, tourism researchers are increasingly applying innovative data-collection methods and sources such as citizen-science and Big Data (Li et al., 2018a; Hu et al., 2018). These new data sources resolve some of the limitations of conventional methods and effectively improve the understanding of tourist behavior.

#### 4.2.2.2. NON-TRADITIONAL DATA SOURCES

As a response to the disadvantages of traditional data sources and inconsistent migration statistics, scholars are continuously searching for alternatives. For instance, several scholars leveraged Internal Revenue Service (IRS) data to study county-to-county and inter-county migrations in the United States (Molloy et al., 2011). Fussell et al. (2014a)

used these data to measure permanent migrations outcomes from Hurricane Katrina, although they acknowledge the inability of this approach to measure temporary displacements. Others have employed USPS mail recipient and vacancy data collected quarterly at the tract level to gauge population movements to and from disaster areas (Finch et al., 2010). While more timely than census and more comprehensive than surveys, USPS data cannot provide individual information about who moved where and when, rather it offers an overall picture of population before, during, and after an event.

In this pursuit of innovative methods, some authors have turned their focus towards passively user-generated geo-referenced data (Goodchild, 2007). Passive citizen-sensor data originate in the data shadow produced by the digital activity of citizens, who leave behind traces of information with multiple potential applications (e.g. advertising, research). Increasingly, many researchers exploit the possibilities of these data in a number of fields such as mobility and transportation (Jurdak et al., 2015), public health (Wesolowski et al., 2012), sociology (Amini et al., 2014) or natural hazards (Li et al., 2018b). To a lesser extent, migration studies have also attempted to exploit this source of information (Zagheni et al., 2014). Migration-related scholars are interested in this data source due to its immediacy (close to real time data), wide coverage, and reduced cost (Spyratos et al., 2018).

Mobile phone call detail records (CDR) hold a tremendous potential for migration studies. However, data accessibility is a large limitation as data are rarely shared by private corporations. Taylor (2016) also debates the ethical and privacy concerns related to this type of data, especially for vulnerable populations in areas of poverty, political instability, or crisis. Although researchers and organizations have used CDR data in mobility studies



(Alexander et al., 2015; Williams et al., 2015), applications in migration are still scarce as it involves massive and long-term datasets. Despite this, some studies demonstrate the potential of phone call data in the field (Bengtsson et al., 2011; Ahas et al., 2018). For instance, Blumenstock (2012) analyzed a 4-year CDR dataset from 1.5 million Rwandans and revealed patterns of temporary and circular migration hidden to surveys conducted by national organizations.

Other sources of Big Data are more accessible to scholars for research purposes and applications abound. State et al. (2013) leveraged repeated logins to Yahoo! to estimate short and medium-term migration flows. Zagheni and Weber (2012) determined age and gender-specific migration rates using a vast sample of Yahoo! e-mail messages. Compared to other data sources, social media emerges as the richest supplier of data for multiple applications (Stock, 2018). Whether exploiting advertising platforms, direct user-generated content (comments, posts, profiles, pictures, etc.), or geo-located information from this user-generated content, social media presents new possibilities for migration research (Laczko, 2015). For instance, Zagheni et al. (2017) developed an innovative application of Facebook's advertising platform to estimate the stock of international migrants in the United States and considered this dataset as a potential continuously updated census. Other social media platforms leveraged in migration studies are Google+ (Messias et al., 2016), LinkedIn (State et al., 2014; Barslund and Busse, 2016), Skype (Kikas et al., 2015), and Twitter (Hawelka et al., 2014; Zagheni et al., 2014). Zagheni et al. (2014) analyzed geolocated tweets from 500,000 users in a 2-year period and concluded that Twitter can be useful to predict turning points in migration trends and to improve the understanding of the relationships between internal and international migration. Hawelka et al. (2014) examined

tweets from 2012 estimating the volume of international travelers by country of residence and identifying spatiotemporal patterns of global mobility.

Social media data for migration research offer both the immediacy and continuous spatiotemporal coverage often lacking in traditional approaches such as surveys and censuses (Spyratos et al., 2018). Considerably large sample sizes and reduced costs are among the most-valued qualities of these data (Zagheni et al., 2018). In a recent report, the European Commission anticipates that Big Data can complement traditional data sources of migration (Hughes et al., 2016). However, scholars must deal with the limitations and weaknesses of these approaches. Selection bias—which relates to any given sample not being representative of the whole population—is one of the main concerns for researchers, as well as privacy and ethical issues (Freudiger et al., 2011; Ruths and Pfeffer, 2014).

The tourism research field is more prolific in the application of innovative data collection methods than migration studies at present, as the characteristics of tourist flows—short-term movements—are more suitable for these approaches. Thus, the literature is rich in examples of active (individuals are aware of the data generation and its purpose) and passive citizen-sensor data approaches such as CDR (Raun et al., 2016), GPS data (Grinberger et al., 2014), Bluetooth data (Versichele et al., 2014), and geo-referenced social media (Girardin et al., 2008; Hawelka et al., 2014). Even post-disaster tourism recovery recently benefited from geotagged social media data applied to study the recovery process after both the magnitude 7.2 Bohol earthquake and super typhoon Haiyan in the Philippines (Yan et al., 2017).

#### 4.3. DISASTER CONTEXT AND STUDY AREA

The 2017 hurricane season was exceptionally active in the Atlantic basin with 17 named storms including six major hurricanes (NOAA, 2017). Three of these major hurricanes — Harvey, Irma, and Maria— made landfall in the United States and are among the top 5 costliest tropical cyclones in the country’s recorded history (NHC, 2018). From September 6 to September 7, 2017, Puerto Rico received the first hurricane impact as Category 5 Hurricane Irma tracked about 60 miles north of the island, far enough to avoid hurricane force winds and a significant storm surge. Even though Puerto Rico did not experience a direct hit from Irma, rainfall totaled 10-15 inches in higher elevations in the central area of the island (Cangialosi et al., 2018). Irma caused three indirect deaths in Puerto Rico as well as widespread power outages, loss of water supply, and minor damage to homes and businesses (Cangialosi et al., 2018). Two weeks later, on September 20, while areas of Puerto Rico were still recovering from Hurricane Irma, Hurricane Maria made landfall along the island’s southeast coast as a category 4 storm. Moving northwestwardly, Maria crossed Puerto Rico leaving a path of complete devastation. The eastern half of the island experienced wind gusts over 200 km/h; some of the most densely populated areas (i.e. San Juan and Carolina) received gusts exceeding 220 km/h (Figure 4.1.A and 4.1.B), which decimated great forest areas knocking down and defoliating numerous trees (Hu and Smith, 2018). The east coast of the island recorded maximum storm surge inundation levels 6 to 9 feet above ground level from the combination of storm surge and the tide (Figure 4.1.A). The storm surge and wave action caused severe damage to buildings, homes, roads, and harbors along the east, southeast, and northeast coast (Pasch et al., 2018). Central portions of the island received more than 25 inches of rainfall from September 19 to September 21

(Figure 4.1.C), with some local stations receiving near 38 inches. River flooding and mudslides were extensive across many parts of the island and caused additional evacuations and rescues in valleys (Pasch et al., 2018).

The effects of Maria in Puerto Rico were catastrophic and triggered a humanitarian crisis that extended several months. The official death toll was considerably underestimated. By December 2017, Puerto Rico's authorities had only recognized 64 direct or indirect deaths (Santiago et al., 2017). The lack of accessibility to remote areas, power and communications outages, as well as the difficulty evaluating indirect deaths from worsening of chronic conditions or from deficiencies in medical treatments, caused a delay in issuing death certificates. A recent study by Kishore et al. (2018) increased the death toll by a factor of 70 to an estimated 4,645 excess deaths in the aftermath of Hurricane Maria (September 20 to December 31).

Even though Puerto Rico's situation was significantly aggravated after Hurricane Maria, the island was already experiencing difficulties long before the 2017 hurricane season. Tracing back to its colonial roots, subsistence agriculture was the most common occupation and underdevelopment, illiteracy, and poverty were rampant on the island for centuries. During the 20th century, under the United States rule, the weak Puerto Rican economy began to diversify and flourish based on favorable federal tax laws. Manufacturing and tourism gained prominence as an important share of Puerto Rico's income. The subsidized economy came to an end in 1996 when President Bill Clinton signed legislation phasing out —over a 10-year period— the favorable tax code that had been active for much of the 20th century. Employment loss followed, after many companies and industries fled the island. The economic model thought to be successful during much of the late 20th and early

21st century had failed in solving the structural problems of poverty, inequality, and dependence (Quiñones-Pérez and Seda-Irizarry, 2016). The Puerto Rican government attempted to recapitalize by issuing a large amount of debt bonds on the eve of the global recession of 2008. The island fell into a debt crisis that exacerbated its employment losses.

Given Puerto Rico's status as a U.S. territory, Puerto Ricans are United States citizens and may travel and migrate freely to the rest of the country. With approximately 45 percent of the population living below the United States federal poverty line and the lack of opportunities on the island, migration towards the mainland United States increased significantly (Quiñones-Pérez and Seda-Irizarry, 2016). In addition, the birth rate continued a decades-long decline and is now among the lowest worldwide. In just a few decades, Puerto Rico's population experienced a shift from a young and rapidly growing population to an aging one where deaths now outnumber births. Between 2005 and 2015, Puerto Rico lost around 400,000 residents (Stone, 2017). The outmigration of hundreds of thousands of skilled professionals and students cast doubt about Puerto Rico's immediate future. In this context, demographic experts anticipate further declines (Stone, 2017), a conclusion concurring with Cross's (2014) expectation that declining populations before a disaster are more likely to experience larger post-disaster population losses.

Along with the demographic forecasts (Figure 4.1.D), health, social, infrastructure stresses, as well as limited government transparency, combine to hinder long-term post-disaster recovery (Government of Puerto Rico, 2018a). The Economic and Disaster Recovery Plan for Puerto Rico recognizes that the island will need deep structural and transformative changes and investments to recover from the existing and systemic

socioeconomic crisis exacerbated by the extensive damage wrought by Hurricane Maria (Government of Puerto Rico, 2018a).

#### 4.4. METHODS

To develop our study, we analyzed over 2.6 billion geotagged tweets comprising a 2-year period from September 1, 2016 to August 31, 2018 on an in-house Big Data computing cluster powered by Hadoop and Impala. The tweets were collected using the Twitter Stream Application Programming Interface (API) with a bounding box covering the whole world. In order to study differences across Puerto Rico, we divided the island in 5 regions – Central, North, West, South, and East (see maps in Figure 4.1).

##### 4.4.1 POST-DISASTER RESIDENTIAL DISPLACEMENT

In order to explore the spatial response of Puerto Rico residents to Hurricane Maria, we first identified local Twitter users by analyzing their tweet activity during the year prior to the impact of Hurricane Irma and Hurricane Maria (September 1, 2016 to August 31, 2017). We assumed those with a majority of tweets originating from Puerto Rico were living on the island. Starting with a collection of tweets sent from Puerto Rico in the pre-disaster period, we followed a three-step process to identify active local users in Puerto Rico (Figure 4.2):

- 1) identification of active users in Puerto Rico during the pre-disaster period,
- 2) retrieval of tweets from identified active users in Puerto Rico for the entire world in the pre-disaster period,

- 3) determination of home location of active users. In this last step, we computed the residential status based on the median center of each user's tweets during the pre-disaster period (Martín et al., 2017).

We then retrieved the tweets from the 32,099 local Puerto Rican users in the entire world during the post-disaster period (September 1, 2017 – August 31, 2018) (Figure 4.2). In order to assure enough temporal accuracy to track the users and characterize the movements of the population after Maria, we reduced these 32,099 users to those who tweeted at least once every month in the post-disaster period. After filtering out accounts with suspicious tweeting behavior (for instance, multi-user Twitter accounts: tweets in distant locations at the same time) and non-human Twitter accounts (sources such as Tweetbot for IS or TweetMyJOBS), the final sample size was composed of 1,231 users. With the purpose of identifying age and gender biases, we visited the user's public profile and examined profile pictures, usernames, full name, description, URLs, multimedia content and tweets uploaded by the users to manually estimate gender (female or male), and approximate age range (17 years or fewer, 18-24, 25-34 years, 35-44 years, 45-54 years, 55-64 years, 65-74 years, and 75 years or older). By individually analyzing the location of tweets from these 1,231 users, we collected information about post-disaster population movements: displacement estimates, destinations, timing, and the return.

We then tested the association of gender, age, region of residence (central region, north region, west region, south region, east region), and residence in a coastal municipality (coastal or non-coastal) with displacement behavior. To measure the association of gender and age (demographics) and location of residence with the displacement outcome, we conducted bivariate Chi-Squared tests of independence. We contextualize our findings and

the suitability of Twitter for assessing post-disaster displacement and migration by comparison with recent studies about these processes in Puerto Rico such as Teralytics (2018), Sutter and Hernandez (2018), Hinojosa et al. (2018), Hinojosa and Meléndez (2018), and United States Census Bureau (2018a).

#### 4.4.2. POST-DISASTER POPULATION INFLOWS

For the analysis of population inflows into Puerto Rico, we compare the amount of non-resident Twitter users active every week during the post-disaster period (September 1, 2017 - August 31, 2018) to baseline pre-disaster levels (September 1, 2016 - August 31, 2017). To distinguish resident users from non-resident Twitter users, we followed an approach similar to that used to study population displacement among Puerto Rico (Figure 4.2). We applied the process to both samples (pre- and post-disaster) and only non-residents were finally counted.

### 4.5. RESULTS AND DISCUSSION

#### 4.5.1. POST-DISASTER RESIDENTIAL DISPLACEMENT

Figure 4.3 presents the comparison between the population pyramid of Puerto Rico from the 2017 American Community Survey (ACS) (United States Census Bureau, 2018b) and the age and gender distribution of the Twitter sample used in this study. The demographic composition of the Twitter users is the first step in recognizing potential biases in our data as well as in contextualizing the findings. The population pyramid confirms that younger age cohorts dominate Twitter and that the representativeness of the Twitter sample for persons over 54 years old decreases considerably. The Twitter sample shows asymmetry in the gender distribution, with a larger presence of females in the age segment 15-24 years and higher percentages of males in older groups. Overall, the sample is composed of 55%



males. This is consistent with studies identifying a general male bias in Twitter samples (Mislove et al., 2011) and with those who find an overrepresentation of females in the youngest cohorts (Leak and Lansley, 2018). With a sample size of 1,231 Twitter users and using the latest ACS population estimate (2017) of 3,337,177 Puerto Rico residents, we calculate a margin of error for our sample of 3.67% with a confidence level of 99%.

Our analysis revealed that 36.4% of the identified Twitter users left Puerto Rico within the 15 weeks (until Dec. 31 2017) after Hurricane Maria made landfall (Table 4.1). Users who traveled outside of Puerto Rico for 4 weeks or less are considered non-displaced. This group (26.5%) consisted of holiday travelers and Puerto Ricans that sought shelter off the island after the hurricane for a shorter duration while the main lifelines (power, water and phone service) were reestablished on the island. Following previous literature (State et al. 2013), we considered as displaced those users whose stays outside of Puerto Rico lasted more than 4 weeks (8.3%).

Numerous reports estimated the number of Puerto Ricans that left the island because of the hurricane. However, estimates vary greatly due to different data-collection methods and analytical approaches. Some of these estimates range from 160,000 (Hinojosa and Meléndez, 2018) to 400,000 people (Teralytics, 2018). Many factors account for the differences and need consideration. For example, we must note that estimates based on Twitter assume all movements are associated with the hurricane which could mask different travel motivations. Also, the Twitter sample is biased towards a more migration-prone population (e.g. younger population) (Abel and Deitz, 2014). Studies regarding the extent of representativeness of the Twitter population in comparison to the overall Puerto Rican population would be needed to quantify all the potential additional biases. Further,

the Teralytics study (2018) did not distinguish between length of stay and include all travels from Puerto Rico to the continental United States, while those studies looking at student population losses reported by Puerto Rico's Department of Education (Hinojosa and Meléndez, 2018) might obscure migrations from older cohorts. For more discussion on different data sources employed in measuring Puerto Rico's post-Maria exodus see Hinojosa and Meléndez (2018).

Table 4.2 compares the destination of the displaced Twitter users with other studies that utilized different collection methods. Although our approach permits county-level destinations, we aggregated the results to the state-level for comparison and validation purposes, as most published studies did not report county-scale information. Florida stands out as the preferred destination for displaced Puerto Ricans (38.8%), followed by New York (7.8%), Massachusetts (7.8%), and Texas (7.8%). This destination pattern is consistent with other studies using call detail records (CDR) (Teralytics, 2018), FEMA change-of-addresses data and FEMA applications for disaster assistance (Hinojosa et al., 2018; Sutter and Hernandez, 2018), U.S. Postal Service change-of-address requests (Sutter and Hernandez, 2018), and school enrollment (Hinojosa et al., 2018). The concentration of the displaced in Florida, New York, Texas, and Massachusetts confirmed studies suggesting that displacement and migration tend to concentrate in areas where the displaced/migrants already have sociocultural ties (McLeman and Hunter, 2010; Findlay, 2011; Herdağdelen et al., 2016). Our results, however, must be interpreted with caution as the destinations in our study are based on a very small sample size (103 displaced).

The majority of the displaced left Puerto Rico in the first half of October (Figure 4.4), likely pushed by the extended duration of power outages. In total, 76.3% of the displaced

left the island within the first six weeks after Maria. This pattern coincides with findings from Hinojosa and Meléndez (2018). The return process was scattered across the following months, with higher rates after holiday periods (Thanksgiving, Christmas) and at the end of the academic year (May). Nine months after the disaster (May 31, 2018), only 54.6% of those displaced had returned to Puerto Rico according to our data. If considering the whole Twitter sample, 3.8% of the Twitter sample had relocated and not returned. As seen earlier, this result should be interpreted in the context of the age biases of the sample. The most up-to-date estimate at the time of this research, based on data released by the United States Census Bureau, shows the displacement of Puerto Ricans is roughly 129,848 people (3.9% of the population) (US Census Bureau, 2018a). This figure closely aligns with our estimates and serves as a relative validation of our approach.

Our results confirmed that Hurricane Maria triggered long-term displacement. However, whether this long-term displacement becomes permanent migration and confirms the most pessimistic population forecasts (Stone, 2017) is still unknown. Drawing causal connections to suggest that severe storms cause permanent migration (Spencer and Urquhart, 2018) is premature at this point.

Table 4.3 presents the results of the chi-squared tests of independence for the variables age, gender, region, and coastal municipality. We found no association between gender ( $\chi^2(1, N=1139) = 1.357, p=0.244$ ) and region ( $\chi^2(4, N=1231) = 2.662, p=0.616$ ) and displacement behavior. The coastal municipality variable shows significant association ( $\chi^2(1, N=1231) = 8.418, p=0.004$ ) but this relationship is very weak ( $\phi = 0.083$ ) and likely caused by the effect size. Age shows a weak (Cramer's  $V = 0.148$ ) but significant association ( $\chi^2(7, N=1081) = 23.524, p=0.001$ ).

Looking more in depth at the results of the relation between age and displacement, the cohort 25-34 is 1.8 times more likely to relocate than the remaining age groups (Table 4.4). This result is statistically significant ( $p < 0.01$ ). The other group with statistically significant results is the 45-54 cohort ( $p < 0.05$ ), which is less likely to relocate outside of Puerto Rico by a factor of 0.1. The loss of the age groups with higher fertility rates, especially in a pre-disaster context of alarmingly low birth rates, can exacerbate the depopulation in the island. This would confirm the expectation that declining populations before a disaster are likely to experience larger post-disaster population losses (Cross, 2014).

#### 4.5.2. TOURISM AND POST-DISASTER POPULATION INFLOWS

The 2017 hurricane season not only accelerated the outmigration of Puerto Ricans towards the continental United States, but it also severely damaged the economy of the island. One of the pillars of this economy, the tourism sector, experienced a major hit. A lack of essential utilities such as power or water, closed airports and cruise terminals, beach erosion, and water contamination were some of the reasons behind a decrease in tourist visitations. Figure 4.5 compares the amount of non-resident Twitter users active in Puerto Rico during the year prior to Hurricane Irma and Hurricane Maria (September 1, 2016 to August 31, 2017) with the non-resident Twitter users active during the post-disaster period (September 1, 2018 to August 31, 2018). The pre-disaster baseline shows a three-peak pattern: 1) the winter break/holiday period (December 15 to January 15) when Puerto Rico attracts tourists due to its warmer climate and beautiful beaches and when many Puerto Ricans residing in the continental United States return home to visit their families; 2) “Spring Break” (February 20 to March 15) when many U. S. college students visit Puerto Rico attracted by the beach and the nightlife of the island; and 3) the summer months (May,

June, July) with visitors especially looking for sun-and-beach activities and food and music festivals in town fairs.

Post-disaster data show that September 2017 began with considerably more non-residents users in the island, but the amount of non-residents fell below 2016-2017 levels shortly after Hurricane Maria (September 20, 2017) (Figure 4.6A). During the first weeks after Hurricane Maria, non-resident user levels stayed close to 2016-2017 totals, likely due to the influx of first responders and relief workers (Government of Puerto Rico, 2018b). The holiday season, beginning with Thanksgiving break and continuing into the winter break revealed a decline of around 30% in the number of non-residents in comparison to pre-disaster levels. The decrease during the spring break high-season had a similar magnitude (30%). In general, looking at Figure 4.6A, we observe larger decreases in high-season periods (winter break, spring break, and summer) than during the low-season. Particularly noticeable is the reduction (around 50%) of non-resident users in the late summer (July 15 to Aug 31) of 2018, which could be related to the negative perception of tourists about the preparedness of Puerto Rico for another active hurricane season in the island (D'Ambrosio, 2018).

The effects of Hurricane Maria on the number of non-resident Twitter users (visitors) were not homogeneously distributed across the island. The hardest hit regions (North and East) (see Figure 4.1→) experienced the largest decreases, while the Central, West, and South had more contained losses or even positive annual balances. The Central region (Figure 4.6B), after a significant increase (over 200%) of non-resident users during September and the first half of October, experienced a slight decrease (5%) until August 2018. The North region received 22% fewer non-resident Twitter users (Figure 4.6C) from

October 2017 to August 2018, with peaks over 30% decrease during high-season periods and over 60% in the late summer. In the same period, the West region (Figure 4.6D) suffered 15% net loss in the total of non-residents Twitter users. However, looking at the intra-annual distribution of this region, we observe two periods where the number of non-residents increased in comparison to baseline levels. First, in the following weeks after Hurricane Maria (from mid-October to mid-December), the totals of non-resident Twitter users increased by 5%. Second, during the spring break high-season period (mid-March), the region experienced a significant increase (11%). The South region (Figure 4.6E) is the only region that saw an increase (9%) in non-resident Twitter users from October 2017 to August 2018. Lastly, the East region (Figure 4.6F), the most severely affected by Hurricane Maria and the most tourism-oriented region, lost 33% non-resident users when comparing to the same period before the hurricane. Here, reductions during high-season periods and the late summer were also more significant than during low-season weeks.

The increase in the number of non-resident Twitter users in the least affected regions (particularly relevant in the South) during the first months after Hurricane Maria –from October to January– reveals that the likely cause was the influx of first responders and relief workers. Although high-season periods were more severely affected throughout in Puerto Rico, we observe how the least affected regions suffered smaller losses, which is probably partly explained by the transfer of visitors from more affected areas (e.g. East region towards West region during the 2018 Spring Break).

#### 4.6. CONCLUSIONS AND FURTHER RESEARCH

The study of population movements, especially during and after disasters, continues to be a major endeavor for authorities and researchers. Traditional sources of data for

migration and tourism studies are often inadequate for estimating the spatiotemporal dimension of the processes in a post-disaster context, especially concerning data accessibility and reliability. Calls for new data sources and approaches abound (Piguet, 2010; Bilborrow and Henry, 2012).

The results presented here confirm the potential for using passive citizen sensor data (Twitter) to estimate the magnitude, timing, destination, and return of the displacement, as well as the extent of the impact on the arrival of non-residents to the island. The findings of this study are consistent with early reports on post-Maria displacement/migration. They also confirm that the hurricane resulted in 8.3% off-island displacement, with nearly 4% of our Twitter sample (mainly composed of 15 to 54 years old individuals) leaving Puerto Rico and not returning as of May 31, 2018. This finding is consistent with Cross (2014), who suggested that places with declining populations before a disaster were likely to experience larger post-disaster population losses. In terms of destinations, 62% of those who relocated for more than 4 weeks moved to Florida, New York, Texas, and Massachusetts, and the timing of departure was concentrated within the first six weeks after Maria (76% of the displaced). Among the variables tested for association with displacement behavior, only age showed a weak association. The age cohort 25-34 years old had a significant and positive association with displacement, corroborating the scenario of the loss of young professionals at their prime fertile age and casting additional doubt over the demographic future of Puerto Rico. Regarding the study of post-disaster population inflows, we can conclude that, as of August 31, 2018, Puerto Rico had not recovered pre-disaster levels of non-resident visitors. However, the geographic patterns

were dissimilar, with the most storm-affected areas (North and East) experiencing larger losses of non-resident visitors than less storm-affected regions (West and South).

Our results indicate that an approach based on geotagged Twitter data is amenable for addressing longstanding problems of data availability and reliability in the displacement/migration and tourism research fields, and in particular for the characterization of the disruption of population fluxes triggered by a disaster. Future research following this line of work might develop close-to-real time information that facilitates the creation of early warning systems for forced displacement following disasters, as well as helping in the monitoring of tourism fluxes which are particularly sensitive to external stressors and elusive of current data-collection methods. Additional research also is required to better understand the representativeness of Twitter data, as whether Twitter users can adequately represent a population is still debatable (Spence et al., 2016).

#### 4.7. ACKNOWLEDGEMENTS

We would like to express our gratitude to Teralytics for sharing the results of their study for incorporation in our research.



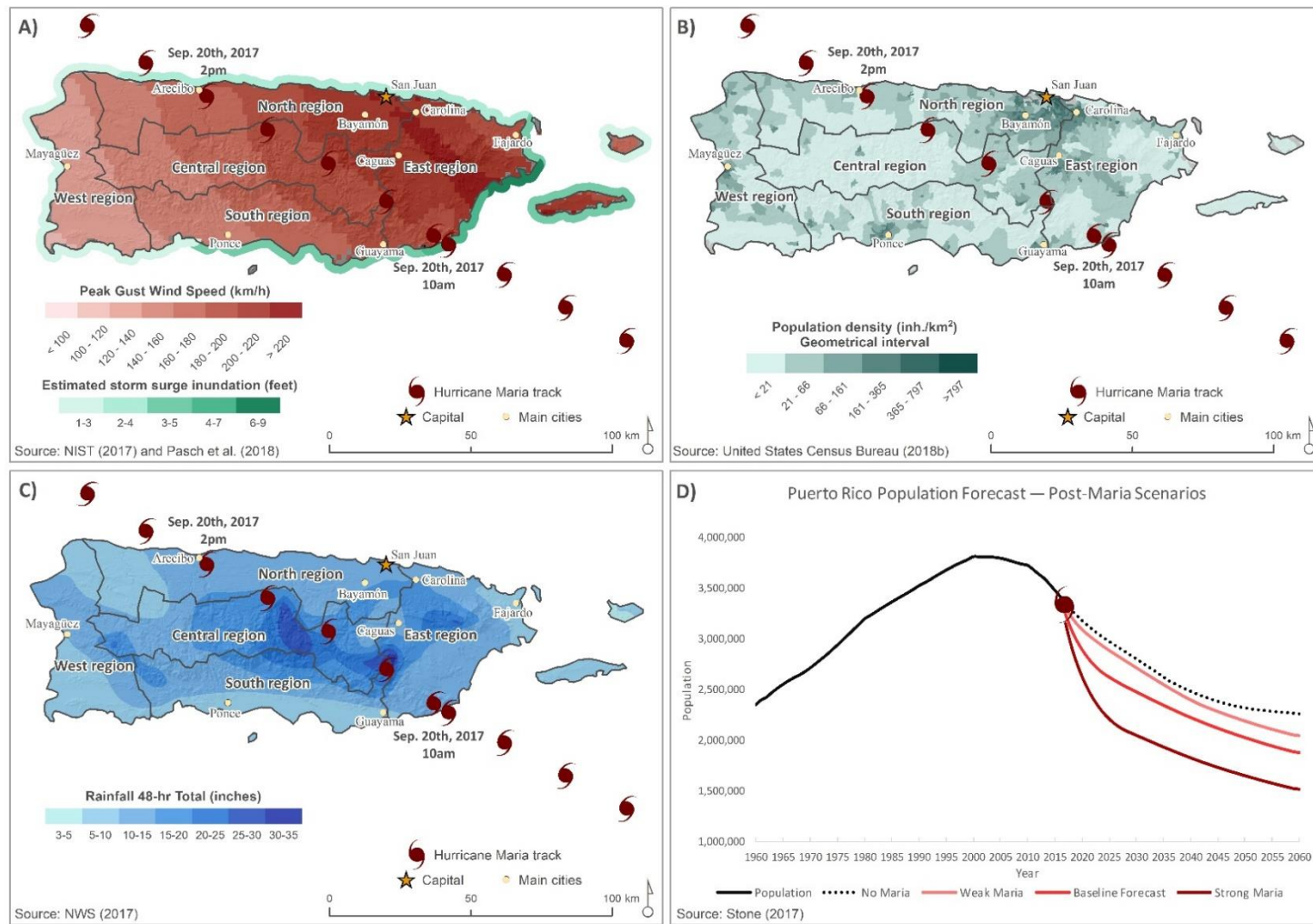


Figure 4.1. Hurricane Maria in Puerto Rico and population forecasts

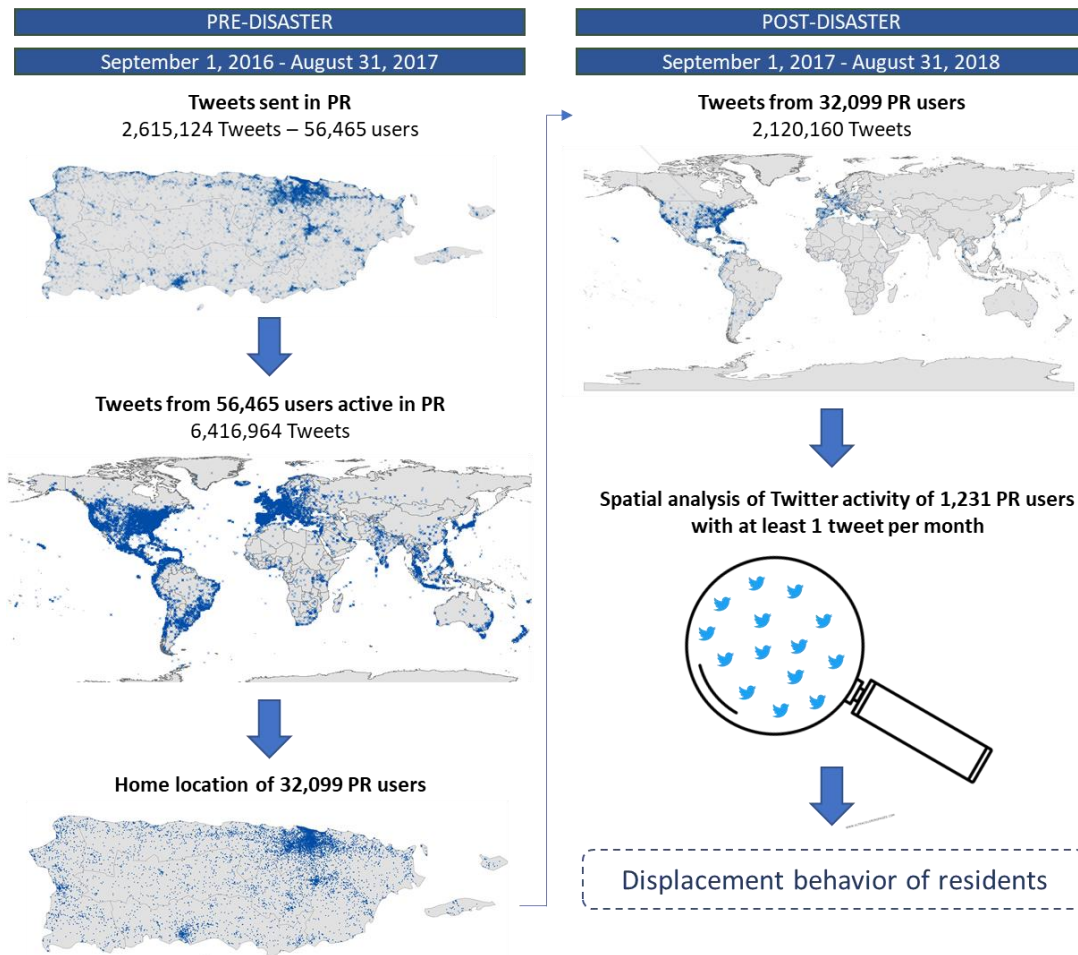


Figure 4.2. Workflow to obtain displacement behavior of Puerto Rican residents

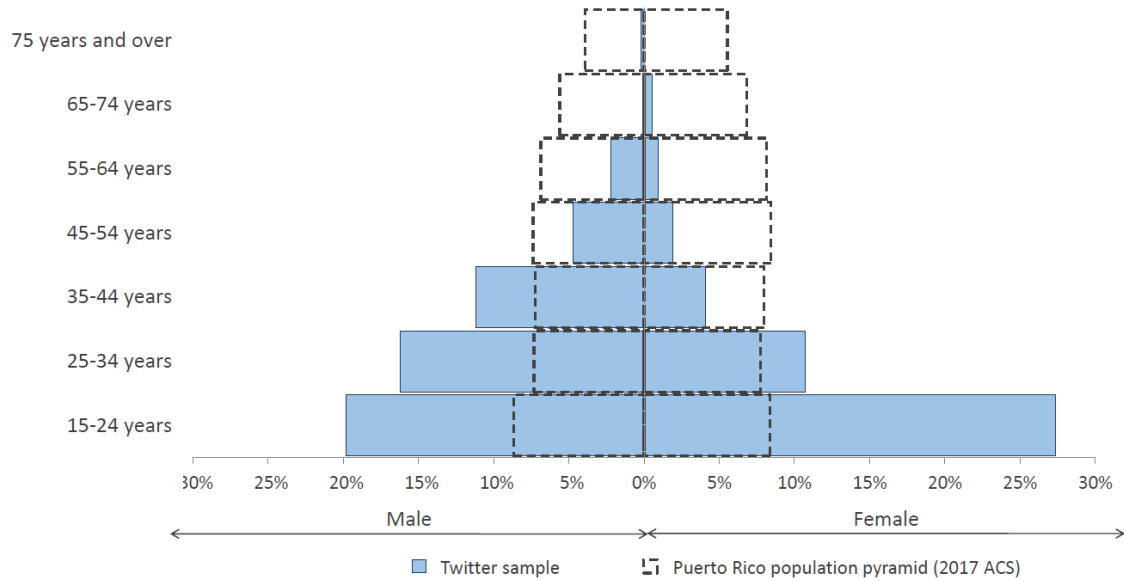


Figure 4.3. Population pyramid of Puerto Rico and Twitter sample.

Table 4.1. Travel duration

	<b>Users</b>	<b>Percentage</b>
Stayed in PR	783	63.6%
Traveled out PR	448	36.4%
4 weeks or less	326	26.5%
5-12 weeks	24	1.9%
13-24 weeks	24	1.9%
24 weeks or more	55	4.5%
Unknown duration	19	1.5%

Table 4.2. Destination of displaced

			TERALYTICS (2018)		SUTTER AND HERNANDEZ (2018)				HINOJOSA ET AL (2018)			
TWITTER			CDR		FEMA		US Postal Service		School Enrollment		FEMA	
TOTAL	%		TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%	TOTAL	%
CALIFORNIA	1	1.0%	6,016	1.7%	159	1.5%	N/A	N/A	N/A	N/A	419	1.0%
CONNECTICUT	2	1.9%	12,339	3.5%	323	3.0%	235	3.5%	1,827	8.2%	2,281	5.7%
FLORIDA	40	38.8%	144,726	41.4%	5,500	51.7%	2,800	42.2%	11,554	51.9%	18,013	45.0%
ILLINOIS	1	1.0%	7,817	2.2%	267	2.5%	103	1.6%	607	2.7%	1,324	3.3%
MASSACHUSETTS	8	7.8%	20,042	5.7%	527	5.0%	345	5.2%	2,556	11.5%	3,399	8.5%
NEW JERSEY	1	1.0%	20,759	5.9%	486	4.6%	268	4.0%	886	4.0%	1,690	4.2%
NEW YORK	8	7.8%	39,060	11.2%	1,003	9.4%	205	3.1%	2,218	10.0%	3,683	9.2%
PENNSYLVANIA	3	2.9%	22,989	6.6%	557	5.2%	874	13.2%	2,599	11.7%	2,954	7.4%
TEXAS	8	7.8%	21,973	6.3%	418	3.9%	393	5.9%	N/A	N/A	1,361	3.4%
VIRGINIA	5	4.9%	5,610	1.6%	156	1.5%	425	6.4%	N/A	N/A	N/A	N/A
TOTAL	77 <sup>a</sup>	74.8%	301,331	86.1%	9,396	88.3%	5,648	85.1%	22,247	100.0%	35,124	87.7%

<sup>a</sup> The total number of displaced is 103, 77 only includes those whose destination was the 10 states shown in this table. Percentages are calculated based on the total number of displaced (103)

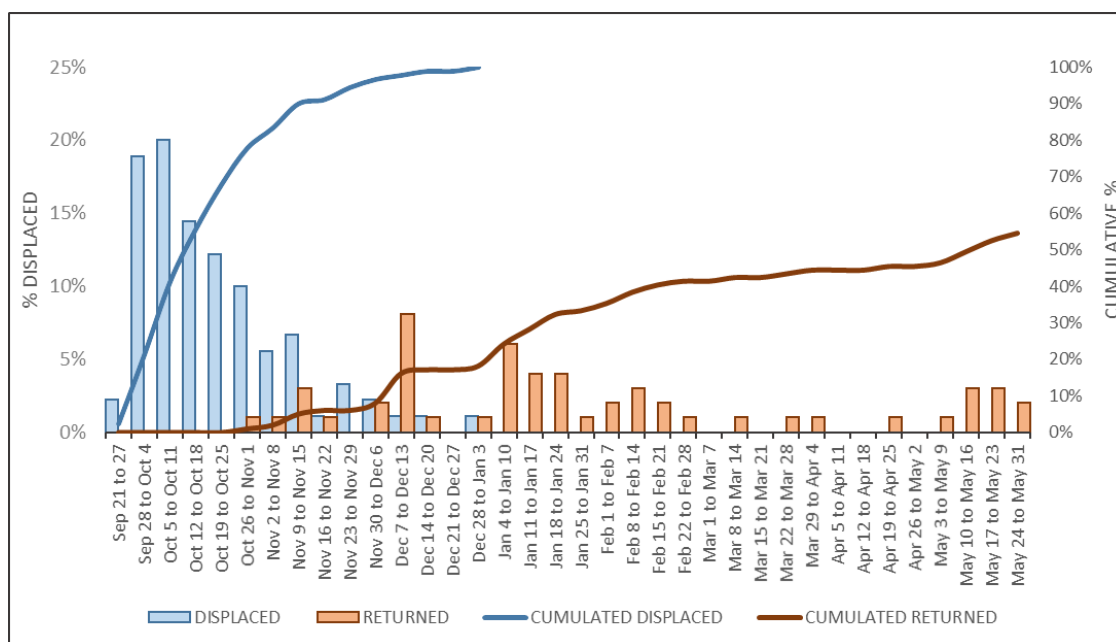


Figure 4.4. Timing of departure and return of the displaced

Table 4.3. Chi-Squared tests of independence

Variable	N	$\chi^2$	p-value	Phi/Cramer's V	df
Age	1081	23.524	0.001**	0.148 <sup>#</sup>	7
Gender	1139	1.357	0.244	-0.035	1
Region	1231	2.662	0.616	0.047 <sup>#</sup>	4
Coastal municipality	1231	8.418	0.004**	0.083	1

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

<sup>#</sup> Cramer's V

Table 4.4. Displacement rates and odds ratio for different age groups

	N	Displaced	Displaced rate	Odds ratio	p-value
<b>Age (ref=remaining groups)</b>	1081				
17 years or fewer	79	3	3.80%	0.4	0.079
18-24 years	431	45	10.40%	1.2	0.313
25-34 years	291	39	13.40%	1.8	0.005**
35-44 years	165	11	6.70%	0.7	0.199
45-54 years	72	1	1.40%	0.1	0.016*
55-64 years	34	0	0.00%	0	0.057
65 years or older	9	2	22.20%	2.8	0.182

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ ,  
\*  $p < 0.05$

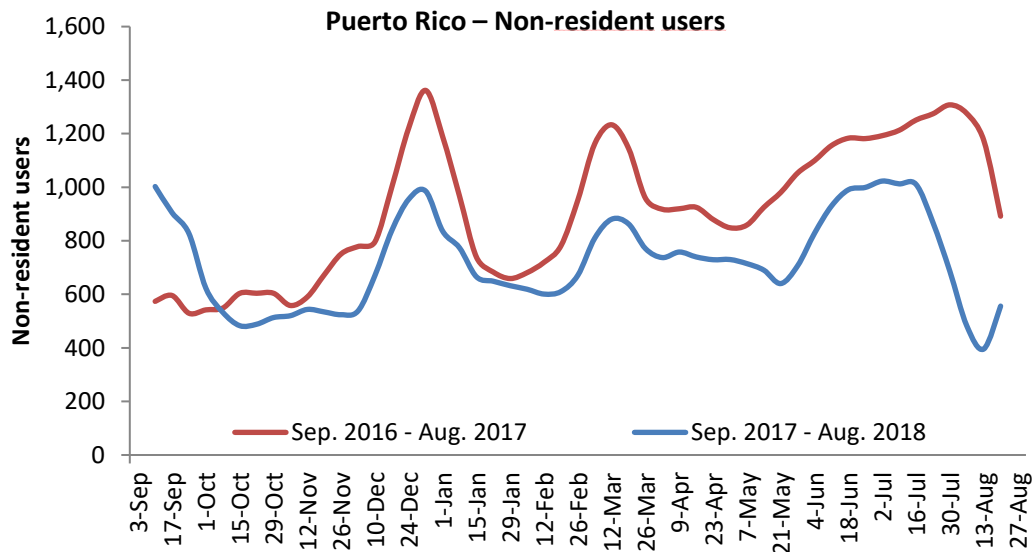


Figure 4.5. 3-week moving average of non-residents Twitter users active in Puerto Rico during the periods Sep. 2016 - Aug. 2017 (pre-disaster) and Sep. 2017-Aug.2018 (post-disaster).



Figure 4.6. Percentage difference of non-resident Twitter users active in Puerto Rico during Sep. 1, 2017 to Aug. 31, 2018 from baseline levels (Sep. 1 2016 to Aug. 31 2017). Green represents positive changes. Red shows negative change.

## **CHAPTER 5**

## **CONCLUSIONS**



### 5.1. TRACKING GEOSPATIAL DIGITAL FOOTPRINTS: AN OPPORTUNITY

The study of population movements has always been a challenge for researchers, particularly limited by the lack of consistent and reliable data regarding both short-term and long-term movements (Spyratos et al., 2018). The development of modern transportation modes facilitated an acceleration in the rate of human motion and enabled the connection of distant places in short periods of time, which added another layer of complexity to the study of population movements. In contexts of increased mobility, many have stressed the inability of traditional methods such as official registries (censuses) and surveys to account for such dynamic processes and urged the development of innovative alternatives and/or supplementary methods (Baker, 2009).

Population movements induced by environmental changes such as disasters are the focus of multiple researchers, particularly after landmark events such as Hurricane Katrina in 2005 (Fussell, 2005). However, studying population movements triggered by hazards or in response to environmental disasters is considered an even more difficult task than regular human mobility in non-stressed situations (Laczko, 2015; Rango and Vespe, 2017). One of the main issues relates to the rapid sequence of population movements during disasters, as these often escape the temporally sparse measures of official population statistics -annual in the best cases. Responding to basic questions like “how many”, “where”, “when”, “who”, or “why” then becomes an extraordinary challenge for which social scientists still try to find optimum solutions.

The advent of digital technologies that has deeply changed society in the last decades is seen as an opportunity in the field of spatial behavior and several scholars are actively exploring the exploitation of geospatial digital trace data for research purposes (Isaacson

and Shoval, 2006). Social scientists agree on the potential value of these digital data shadows generated in much of the digital daily lives of people, however many questions regarding its validity, representativeness, and application still remain unformulated or unanswered. Although the application of approaches based on passive citizen sensor data is increasing in many disciplines, the hazards/disaster field has not fully explored their use for the study of population movements and its association with hazards and emergencies. Only a few studies, mostly conducted by epidemiologists, have systematically leveraged passive citizen sensor data (call data records) to improve the understanding of population movements and its relationship with infectious diseases.

Responding to calls for innovation in techniques for the field of spatial behavior and motivated by the lack of in-depth studies about the suitability of geospatial digital trace data for the study of hazard-triggered population movements, this dissertation extensively explored geotagged tweets as a proxy to expand the understanding of the spatial responses of populations during hurricanes. The results of this research confirm the potential of geotagged social media to complement traditional methods of spatial behavior, recognizing at the same time the limitations of the approach.

#### 5.1.1. HURRICANE-TRIGGERED POPULATION MOVEMENTS: THE TWITTER GAUGE

The dissertation reveals that geotagged tweets are particularly suitable for tracking individuals during the response and recovery phases of a hurricane. Thus, the Twitter-based approach permitted relatively accurate estimations of the magnitude (“how many”), temporal sequence (“when”), and destination (“where”) of inter-county evacuation and

post-disaster displacement processes, as well as estimations of the number of incoming visitors to disaster-stricken areas.

The case study of the evacuation of Hurricane Matthew in South Carolina (Chapter 2 and Chapter 3) demonstrated that geotagged tweets can offer valuable information to researchers and emergency managers. The analysis showed that the temporal scale of a hurricane evacuation (hours to days) is appropriate for the application of an approach based on geotagged social media, as the temporal resolution (periodicity) of geotagged tweets is sufficient to track the location of individuals. Conversely, the spatial resolution of geotagged tweets requires a more detailed discussion. Tweets can be pinpointed to a location in different ways: point (coordinates), neighborhood, city, state, or country. However, only a small fraction of these tweets has a resolution finer than city-level. Because of this, and to have a representative sample in numeric terms, the maximum spatial resolution that the proposed Twitter-based approach can offer is county level aggregations. This spatial resolution is enough to gauge inter-county movement during evacuations, but it cannot properly estimate intra-county movements. The Twitter-based approach determines the evacuation status of the residents based on their location inside or outside a county where a mandatory evacuation is issued. Evacuation zones are very variable in terms of extension as their design is based on a combination of local geographical, infrastructural, and regulatory factors. In some states, like South Carolina, hurricane evacuation zones often enclose whole counties, forcing the population to seek shelter away from their home county. However, in Florida, evacuation zones are narrower along the coastline and all-county evacuations are rare, which enable some of the residents of evacuation zones to evacuate to shelters or houses of relatives/friends within their own

county. The proposed method is therefore valid for areas where all-county evacuations are a common procedure, but it must be applied with caution where evacuation zones do not completely correspond with county limits, as the method might obscure intra-county evacuations and reflect misleading evacuation compliance estimates.

The dissertation also demonstrated that geotagged tweets are also a valid source of information for understanding the population dynamics after a hurricane has passed. Focusing on Hurricane Maria in Puerto Rico as a study case, the Twitter-based approach obtained estimations that aligned with early reports on displacement by the United States Census Bureau. In this case, the temporal scale of the displacement process (weeks/months) is also appropriate for the application of an approach based on geotagged tweets. Chapter 4 demonstrates how Twitter can provide a large-enough sample of active users (at least a geotagged tweet every month) whose tweeting activity is sufficient to determine the displacement timing, the return process (if any), and the destination. In addition, as hurricanes can deeply alter the environment of the affected areas and therefore the livelihood of its residents, Chapter 4 also explored the response of visitors (non-residents) in a disaster-stricken area. The method managed to estimate the reduction in the number of non-resident Twitter users during the months following Hurricane Maria, as well as identifying significant spatiotemporal differences across Puerto Rico.

Although geotagged tweets are considered a spatially sparse and temporally episodic source of data (Andrienko et al., 2012), this research demonstrates that this data source is suitable for gauging population movements during hurricanes. The spatiotemporal scale of the event constitutes the critical aspect to determine the potential application of geotagged tweets for the study of associated population movements. Based on the positive results, the

proposed method could also benefit the study of population spatial response during other large-scale hazards such as earthquakes, tsunamis, volcanoes, droughts, or major floods. Conversely, smaller-scale events such as tornadoes or terrorist attacks are beyond the applicability of a Twitter-based approach and would require finer-resolution passive citizen sensor data such as call data records.

#### 5.1.2 SPATIAL BEHAVIOR IN HURRICANES: THE TWITTER DEPTH

Sui and Goodchild (2011) discussed how the advent of digital technologies and the emergence of geospatial digital trace data brought us closer to knowing where everybody and everything is located at any time. In this context, they also questioned the appropriateness of Manovich's division of social sciences in disciplines based on "surface data about the many" (e.g. sociology, economics, political science, or geography) and disciplines based on "deep data about the few" (e.g. psychology, psychoanalysis, or anthropology) (Manovich, 2011). Sui and Goodchild (2011) reformulated this idea and defended that, with digital geospatial trace data, "deep data about the many" is now possible.

Scholars researching spatial behavior in emergency situations are particularly interested in the differential spatial response of subpopulations, and they therefore look for answers to questions such as "who", "why", and "how". Regarding the first of these questions, public access to the profile of Twitter users permits an estimation of characteristics of the users such as gender, age, and race. Meanwhile, the study of the spatial tweeting behavior allows a resident/non-resident status classification. Thus, Chapter 3 demonstrated that geotagged tweets can further our understanding of understudied subpopulations such as the young, racial minorities, or non-residents during evacuations.

These have been groups traditionally underrepresented in conventional approaches such as survey questionnaires and where a Twitter-based approach can be of use to complement these analyses. However, the range of demographic characteristics that can be retrieved from Twitter is limited (in number and accuracy) and based on inferences (low confidence), which is in contrast of self-reported (high confidence) information from surveys. The validity of these inferences needs further investigation, as well as the representativeness of Twitter users who share their geotagged tweets regarding the overall population.

The second major question (“why”) constitutes the major limitation of the Twitter-based approach. Again, the association of the spatial movement of a Twitter user with an upcoming threat or the post-disaster environmental disruption is founded on inferences about their motivation. In the case of evacuations (Chapter 2 and Chapter 3), the Twitter-based approach assumed that any inter-county movement between the pre-evacuation phase and the post-evacuation phase was directly associated with the hurricane evacuation. Even though it is logical that most of these movements were indeed associated with the evacuation process, the Twitter-based approach cannot offer any kind of validation that allows confirming the motivation of the trips. This fact means that within the group of Twitter users considered as evacuees the approach could include users whose movement motivation was work-related or holiday-related. In geotagged tweet-based studies the reasoning behind evacuating or staying in place also remains unknown, which limits its actual direct application for the understanding of critical factors such as risk and uncertainty of motility. The method encounters the same limitation in relation to post-disaster displacement (Chapter 4), as it cannot provide certainty about the reason of people’s movements. An avenue to overcome this limitation would imply the use of

“excess mobility analyses”. Similar to the concept of “excess mortality”, largely used in the hazard/disaster field to account for indirect deaths of a disaster (e.g. heat and cold waves, tropical cyclones etc.) when the cause of death is not easily connected to the event, a statistical analysis of Twitter “excess mobility” could shed light into the accuracy of the population movement estimates of this dissertation. This idea was beyond the scope of the dissertation. However, provided that long-term longitudinal passive citizen sensor data are available, it could offer a new way of looking at population movements.

Answers to the “how” question, and many other aspects related to the evacuation or displacement processes that researchers are able to ask during a survey or an interview, are impossible to acquire by exploiting digital trace data. This brings us back to the idea of “deep data about the many” (Sui and Goodchild, 2011). When referring to passive citizen sensor data in emergency situations, the topic of this dissertation, the “depth” of geotagged tweets is insufficient to explore many aspects of evacuation and displacement behavior, although the results of this dissertation have demonstrated a certain degree of application. As of now, geotagged tweets cannot be considered deep data. Other passive citizen sensor data such as call data records might facilitate a more detailed background of the cell phone user. However, cell phone carriers cannot share individual information due to privacy regulations. Aggregates of this information have been occasionally shared in the past for research purposes. These could be of help to understand more about spatial behavior in stressed situations and test new hypotheses, however, accessibility remains severely restricted. “Deep data for the many” is a concept still far from becoming a reality in the field.

## 5.2. A GLANCE AT THE FUTURE

Isaacson and Shoval (2006) claimed that the study of human behavior was entering a new phase due to the possibility of spatially tracking individuals through their digital footprints, and thus being able to keep up with the spatiotemporal acceleration of human life. More than a decade later, this dissertation is an example of the potential application of passive citizen sensor data for application in the hazards/disaster field and particularly in the estimation of emergency-triggered population movements and certain aspects of spatial behavior under-stressed situations. This work explored the most accessible source of geospatial digital trace data currently available, therefore serving as a first step for a data source audit that reveals the opportunities and obstacles for the advancement of the field and building the foundation of this new phase in behavioral geography.

This research therefore opens new possibilities for complementing traditional approaches for the study of human spatial behavior. Passive citizen sensor data cannot by any means substitute the information richness that a well-designed survey questionnaire can provide, but it can certainly help overcome some of the long-known limitations of these traditional approaches. For instance, one of the most relevant characteristics of passive citizen sensor data, and more specifically geotagged social media, is its velocity. The analyses performed in this research have the potential to be automated and performed in near real time, allowing an accurate situational analysis of any emergency-triggered population movement. Advancing in this direction, passive citizen sensor data can produce a dynamic census that helps emergency managers estimate evacuation compliance or detect early signs of post-disaster displacement, which could be a great improvement particularly in developing countries, where cell phone and Internet penetration is considerably growing



but authorities often have no capacity of estimating population movement, and even less so in disaster situations. This information can also form part of models that help us predict what the future might hold in terms of post-disaster displacement and thus assist in one of the least understood aspects of the disaster cycle, the recovery phase.

The main obstacle in the advance towards dynamic population statistics through the exploitation of geospatial digital trace data is data accessibility. The great majority of these data are privately owned (e.g. mobile phone operators, providers of social media platforms or other internet-based services) and most companies refuse to disclose these data for research purposes because of the importance of these data for their market competitiveness. Only a few companies, such as Twitter, publicly share a small part of their data and permit its use for different projects. Other companies have occasionally facilitated data-sharing for research purposes, however this practice is not extensive. As discussed in this dissertation, the small fraction of geotagged tweets is not the richest form of passive citizen sensor data available, and a number of projects that leverage this data source is not related to its quality or validity but to the possibility of accessing the data. The “new phase” for spatial behavior is at risk of not materializing if social scientists continue to be kept out of the data. If we decide to advance in this path, a discussion is needed on a new public-private partnership framework that deals with ethics and privacy concerns and that guarantees a responsible use of geospatial digital trace data. Until then, researchers will be severely limited, and our research will only be able to show what the study of spatial behavior could be like in this new era.

## REFERENCES

- Abel, J., & Deitz, R. (2014). The causes and consequences of Puerto Rico's declining population. *Current Issues in Economics and Finance*, 20(4), 1-8. Retrieved from [https://www.newyorkfed.org/medialibrary/media/research/current\\_issues/ci20-4.pdf](https://www.newyorkfed.org/medialibrary/media/research/current_issues/ci20-4.pdf)
- Acredolo, L. P., & Evans, D. (1980). Developmental changes in the effects of landmarks on infant spatial behavior. *Developmental Psychology*, 16(4), 312–318. doi:10.1037/0012-1649.16.4.312
- Adey, P. (2016). Emergency mobilities. *Mobilities*, 11(1), 32-48. doi:10.1080/17450101.2015.1103533
- Adger, W. N., de Campos, R. S., & Mortreux, C. (2018). Mobility, displacement and migration, and their interactions with vulnerability and adaptation to environmental risks. In McLeman, R., & Gemenne, F. (Eds.). *Routledge Handbook of Environmental Displacement and Migration* (pp. 29-41). New York: Routledge. doi:10.4324/9781315638843-3
- Ahas, R., Silm, S., & Tiru, M. (2018, January). Measuring transnational migration with roaming datasets. In *Adjunct Proceedings of the 14th International Conference on Location Based Services* (pp. 105-108). ETH Zurich.
- Alexander, L., Jiang, S., Murga, M., & González, M. C. (2015). Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transportation Research Part C: Emerging Technologies*, 58, 240-250. doi:10.1016/j.trc.2015.02.018
- Amini, A., Kung, K., Kang, C., Sobolevsky, S., & Ratti, C. (2014). The impact of social segregation on human mobility in developing and industrialized regions. *EPJ Data Science*, 3(1), 1–20. doi:10.1140/epjds31

- Anderson, J. (1971). Space-time budgets and activity studies in urban geography and planning. *Environment and Planning A*, 3(4), 353–68. doi:10.1068/a030353
- Andrienko, N., Andrienko, G., Stange, H., Liebig, T., & Hecker, D. (2012). Visual analytics for understanding spatial situations from episodic movement data. *KI-Künstliche Intelligenz*, 26(3), 241-251. doi:10.1007/s13218-012-0177-4
- Archibald, E., & McNeil, S. (2012). Learning from traffic data collected before, during and after a hurricane. *IATSS Research*, 36(1), 1-10. doi:10.1016/j.iatssr.2012.06.002
- Argent, N. (2017). Behavioral geography. *International Encyclopedia of Geography: People, the Earth, Environment and Technology*: 1-11. doi:10.1002/9781118786352.wbieg0875
- Arlikatti, S., Lindell, M. K., Prater, C. S., & Zhang, Y. (2006). Risk area accuracy and hurricane evacuation expectations of coastal residents. *Environment and Behavior*, 38(2), 226–247. doi:10.1177/0013916505277603.
- Bajardi, P., Delfino, M., Panisson, A., Petri, G., & Tizzoni, M. (2015). Unveiling patterns of international communities in a global city using mobile phone data. *EPJ Data Science*, 4(3), 1-17. doi:10.1140/epjds/s13688-015-0041-5.
- Baker, E. J. (1991). Hurricane evacuation behavior. *International Journal of Mass Emergencies and Disasters*, 9(2), 287-310.
- Baker, E. J. (1997). *Hurricanes Bertha and Fran in North and South Carolina: Evacuation behavior and attitudes toward mitigation* (Report prepared for the Charleston and Wilmington districts). US Army Corps of Engineers. Tallahassee, FL: Hazards Management Group, Inc.
- Baker, E. J. (2000). Hurricane evacuation in the United States. In: Pielke RA. Jr., & Pielke RA (Eds.). *Storms* (pp. 306-319) London: Routledge.
- Baker, E. J. (2009). Acquiring data on travel behaviour during emergencies and exceptional events. In Bonnel, P, Lee-Gosselin, M., Zmud, J., & Madre, J.L. (Eds.). *Transport Survey Methods: Keeping up with a Changing World* (pp. 197–211). Bingley: Emerald Group Publishing Limited. doi:10.1108/9781848558458-011.

- Ballard, A. J., Ullman, B. R., Trout, N. D., Venglar, S. P., Borchardt, D. W., Voigt, A. P., ... & Rajbhandari, R. (2008). *Hurricane evacuation traffic operations* (No. FHWA/TX-08/0-4962-1). Retrieved from <https://static.tti.tamu.edu/tti.tamu.edu/documents/0-4962-1.pdf>
- Baro, M., & Deubel, T. F. (2006). Persistent hunger: Perspectives on vulnerability, famine, and food security in sub-Saharan Africa. *Annual Review of Anthropology*, 35, 521-538.  
doi:10.1146/annurev.anthro.35.081705.123224
- Barslund, M., & Busse, M. (2016). *How mobile is tech talent? A case study of IT professionals based on data from LinkedIn* (CEPS Special Report, No. 140). Retrieved from <https://ssrn.com/abstract=2859399>
- Bateman, J. M., & Edwards, B. (2002). Gender and evacuation: A closer look at why women are more likely to evacuate for hurricanes. *Natural Hazards Review*, 3(3), 107-117.  
doi:10.1061/(ASCE)1527-6988(2002)3:3(107)
- Bauman, Z. (1996). From pilgrim to tourist—or a short history of identity. *Questions of Cultural Identity*, 1, 18-36.  
doi:10.4135/9781446221907.n2
- Bauman, Z. (2000). *Liquid modernity*. Cambridge: Polity Press
- Bell, M., Charles-Edwards, E., Kupiszewska, D., Kupiszewski, M., Stillwell, J., & Zhu, Y. (2015). Internal migration data around the world: Assessing contemporary practice. *Population, Space and Place*, 21(1), 1-17.  
doi:10.1002/psp.1848
- Bengtsson, L., Lu, X., Thorson, A., Garfield, R., & Von Schreeb, J. (2011). Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: A post-earthquake geospatial study in Haiti. *PLoS Medicine*, 8(8), e1001083.  
doi:10.1371/journal.pmed.1001083
- Bilsborrow, R. E. (2002). Migration, population change, and the rural environment. *Environmental Change and Security Project Report*, 8(1), 69-84. Retrieved from [https://www.wilsoncenter.org/sites/default/files/Report\\_8\\_Bilsborrow\\_article.pdf](https://www.wilsoncenter.org/sites/default/files/Report_8_Bilsborrow_article.pdf)

- Bilsborrow, R. E., & Henry, S. J. (2012). The use of survey data to study migration–environment relationships in developing countries: Alternative approaches to data collection. *Population and Environment*, 34(1), 113-141. doi:10.1007/s11111-012-0177-1
- Birenboim, A., & Shoval, N. (2016). Mobility research in the age of the smartphone. *Annals of the American Association of Geographers*, 106(2), 283-291. doi:10.1080/00045608.2015.1100058
- Black, R., Arnell, N. W., Adger, W. N., Thomas, D., & Geddes, A. (2013). Migration, immobility and displacement outcomes following extreme events. *Environmental Science & Policy*, 27, S32-S43. doi:10.1016/j.envsci.2012.09.001
- Blumenstock, J. E. (2012). Inferring patterns of internal migration from mobile phone call records: Evidence from Rwanda. *Information Technology for Development*, 18(2), 107-125. doi:10.1080/02681102.2011.643209
- Bonham, C., Edmonds, C., & Mak, J. (2006). The impact of 9/11 and other terrible global events on tourism in the United States and Hawaii. *Journal of Travel Research*, 45(1), 99-110. doi:10.1177/0047287506288812
- Bowser, G. C. (2013). *Determining the differences in hurricane perception and evacuation behavior in the elderly of South Carolina* (Doctoral dissertation). University of South Carolina, Columbia, SC, USA. Retrieved from <https://scholarcommons.sc.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=3481&context=etd>
- Bowser, G. C., & Cutter, S. L. (2015). Stay or go? Examining decision making and behavior in hurricane evacuations. *Environment: Science and Policy for Sustainable Development*, 57(6), 28-41. doi:10.1080/00139157.2015.1089145
- Brommer, D. M., & Senkbeil, J. C. (2010). Pre-landfall evacuee perception of the meteorological hazards associated with Hurricane Gustav. *Natural Hazards*, 55(2), 353-369. doi:10.1007/s11069-010-9532-7
- Bukvic, A., Smith, A., & Zhang, A. (2015). Evaluating drivers of coastal relocation in Hurricane Sandy affected communities. *International Journal of Disaster Risk Reduction*, 13, 215-228.

doi:10.1016/j.ijdr.2015.06.008

- Burke, S., Bethel, J. W., & Britt, A. F. (2012). Assessing disaster preparedness among latino migrant and seasonal farmworkers in eastern North Carolina. *International Journal of Environmental Research and Public Health*, 9(9), 3115–3133. doi:10.3390/ijerph9093115
- Burton, S. H., Tanner, K. W., Giraud-Carrier, C. G., West, J. H., & Barnes, M. D. (2012). "Right time, right place" health communication on Twitter: Value and accuracy of location information. *Journal of Medical Internet Research*, 14(6). doi:10.2196/jmir.2121.
- Cahyanto, I., Pennington-Gray, L., Thapa, B., Srinivasan, S., Villegas, J., Matyas, C., & Kioussis, S. (2014). An empirical evaluation of the determinants of tourist's hurricane evacuation decision making. *Journal of Destination Marketing & Management*, 2(4), 253-265. doi:10.1016/j.jdmm.2013.10.003
- Cangialosi, J. P., Latto, A. S., & Berk, R. (2018). *Hurricane Irma (AL112017)*. Retrieved from National Oceanic and Atmospheric Administration, National Hurricane Center website: [https://www.nhc.noaa.gov/data/tcr/AL112017\\_Irma.pdf](https://www.nhc.noaa.gov/data/tcr/AL112017_Irma.pdf)
- Chae, J., Thom, D., Jang, Y., Kim, S., Ertl, T., & Ebert, D. S. (2014). Public behavior response analysis in disaster events utilizing visual analytics of microblog data. *Computers & Graphics*, 38, 51-60. doi:10.1016/j.cag.2013.10.008
- Chaintreau, A., Hui, P., Crowcroft, J., Diot, C., Gass, R., & Scott, J. (2007). Impact of human mobility on opportunistic forwarding algorithms. *IEEE Transactions on Mobile Computing*, 6, 606–620. doi:10.1109/infocom.2006.172
- Chang, S. E. (2010). Urban disaster recovery: A measurement framework and its application to the 1995 Kobe earthquake. *Disasters*, 34(2), 303-327. doi:10.1111/j.1467-7717.2009.01130.x
- Chen, X., Elmes, G., Ye, X., & Chang, J. (2016). Implementing a real-time Twitter-based system for resource dispatch in disaster management. *GeoJournal*, 81(6), 863-873. doi:10.1007/s10708-016-9745-8

- Çolak, S., Lima, A., & González, M. C. (2016). Understanding congested travel in urban areas. *Nature Communications*, 7, 10793. doi:10.1038/ncomms10793.
- Colleoni, M. (2016). A social science approach to the study of mobility: An introduction. In Pucci, P., & Colleoni, M. (Eds.). *Understanding Mobilities for Designing Contemporary Cities* (pp. 23-33). Cham: Springer. doi:10.1007/978-3-319-22578-4\_2
- Collins, J., Ersing, R., Polen, A., Saunders, M., & Senkbeil, J. (2018). The effects of social connections on evacuation decision making during Hurricane Irma. *Weather, Climate, & Society*, 10(3), 459-469. doi:10.1175/WCAS-D-17-0119.1
- Comerio, M. C. (2005, January). Key elements in a comprehensive theory of disaster recovery. In *1st International Conference on Urban Disaster Reduction*, (Vol. 3). Institute for Social Safety Science and Earthquake Engineering Research Institute (EERI), Kobe, Japan.
- Cooper, M., Biehl, J., Filby, G., & Kratz, S. (2016). LoCo: Boosting for indoor location classification combining Wi-Fi and BLE. *Personal and Ubiquitous Computing*, 20(1), 83-96. doi:10.1007/s00779-015-0899-z.
- Cresswell, T. (2011). Mobilities I: Catching up. *Progress in Human Geography*, 35(4), 550-558. doi:10.1177/0309132510383348
- Crooks, A., Croitoru, A., Stefanidis, A., & Radzikowski, J. (2013). #Earthquake: Twitter as a distributed sensor system. *Transactions in GIS*, 17(1), 124-147. doi:10.1111/j.1467-9671.2012.01359.x
- Cross, J. A. (2014). Disaster devastation of US communities: Long-term demographic consequences. *Environmental Hazards*, 13(1), 73-91. doi:10.1080/17477891.2013.864594
- Crossett, K., Ache, B., Pacheco, P., & Haber, K. (2013). *Population trends from 1970 to 2020* (National Coastal Population Report). Retrieved from <http://oceanservice.noaa.gov/facts/coastal-population-report.pdf>
- Culotta, A., Ravi, N. K., & Cutler, J. (2016). Predicting Twitter user demographics using distant supervision from website traffic data. *Journal of Artificial Intelligence Research*, 55, 389-408. doi:10.1613/jair.4935

- Curtis, K. J., & Bergmans, R. S. (2018). Estimating the population impacts of sea level rise. In McLeman, R., & Gemenne, F. (Eds.). *Routledge Handbook of Environmental Displacement and Migration* (pp. 106-116). New York: Routledge. doi:10.4324/9781315638843-8
- Curtis, K. J., Fussell, E., & DeWaard, J. (2015). Recovery migration after Hurricanes Katrina and Rita: Spatial concentration and intensification in the migration system. *Demography*, 52(4), 1269-1293. doi:10.1007/s13524-015-0400-7
- Cutter, S. L., Emrich, C. T., Bowser, G. C., Angelo, D., & Mitchell, J. T. (2011). *South Carolina hurricane evacuation behavioral study*. Retrieved from [http://webra.cas.sc.edu/hvri/docs/HES\\_2011\\_Final\\_Report.pdf](http://webra.cas.sc.edu/hvri/docs/HES_2011_Final_Report.pdf)
- Cutter, S. L., & Smith, M. M. (2009). Fleeing from the hurricane's wrath: Evacuation and the two Americas. *Environment: Science and Policy for Sustainable Development*, 51(2), 26-36. doi:10.3200/ENVT.51.2.26-36
- Dalakoglou, D., & Harvey, P. (2012). Roads and anthropology: Ethnographic perspectives on space, time and (im) mobility. *Mobilities*, 7(4), 459-465. doi:10.1080/17450101.2012.718426
- D'Ambrosio, R. (2018, May 14). As hurricane season nears, travel agents wonder if Puerto Rico is ready. *Travelmarket*. Retrieved from <https://www.travelmarketreport.com/articles/As-Hurricane-Season-Nears-Travel-Agents-Wonder-if-Puerto-Rico-Is-Ready>
- Dash, N., & Gladwin, H. (2007). Evacuation decision making and behavioral responses: Individual and household. *Natural Hazards Review*, 8(3), 69-77. doi:10.1061/(asce)1527-6988(2007)8:3(69)
- Dash, N., & Morrow, B. H. (2000). Return delays and evacuation order compliance: The case of Hurricane Georges and the Florida Keys. *Global Environmental Change Part B: Environmental Hazards*, 2(3), 119-128. doi:10.1016/S1464-2867(01)00008-0
- Data-Pop Alliance. (2015). *Big Data for climate change and disaster resilience: Realising the benefits for developing countries*. Retrieved from <http://datapopalliance.org/wp-content/uploads/2015/11/Big-Data-for-Resilience-2015-Report.pdf>



- David, C. C., Ong, J. C., & Legara, E. F. T. (2016). Tweeting Supertyphoon Haiyan: Evolving functions of Twitter during and after a disaster event. *PloS ONE*, 11(3), e0150190. doi:10.1371/journal.pone.0150190 PMID: 27019425
- de Albuquerque, J. P., Herfort, B., Brenning, A., & Zipf, A. (2015). A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International Journal of Geographical Information Science*, 29(4), 667-689. doi:10.1080/13658816.2014.996567
- Demuth, J. L., Morss, R. E., Lazo, J. K., & Trumbo, C. (2016). The effects of past hurricane experiences on evacuation intentions through risk perception and efficacy beliefs: A mediation analysis. *Weather, Climate, & Society*, 8(4), 327-344. doi:10.1175/WCAS-D-15-0074.1
- Dixon, D. S., Mozumder, P., Vásquez, W. F., & Gladwin, H. (2017). Heterogeneity Within and Across Households in Hurricane Evacuation Response. *Networks and Spatial Economics*, 17(2), 645-680. doi:10.1007/s11067-017-9339-0
- Dobra, A., Williams, N. E., & Eagle, N. (2015). Spatiotemporal detection of unusual human population behavior using mobile phone data. *PloS ONE*, 10(3), e0120449. doi:10.1371/journal.pone.0120449
- Dong, H., Halem, M., & Zhou, S. (2013, September). Social media data analytics applied to hurricane sandy. In *2013 International Conference on Social Computing* (pp. 963-966). IEEE. Washington, DC, USA doi:10.1109/SocialCom.2013.152
- Dow, K., & Cutter, S. L. (1998). Crying wolf: Repeat responses to hurricane evacuation orders. *Coastal Management*, 26(4), 237-52. doi:10.1080/08920759809362356.
- Dow, K., & Cutter, S. L. (2000). Public orders and personal opinions: Household strategies for hurricane risk assessment. *Global Environmental Change Part B: Environmental Hazards*, 2(4), 143-155. doi:10.1016/S1464-2867(01)00014-6
- Dow, K., & Cutter, S. L. (2002). Emerging hurricane evacuation issues: Hurricane Floyd and South Carolina. *Natural Hazards Review*, 3(1), 12-18. doi:10.1061/(ASCE)1527-6988(2002)3:1(12)
- Drabek, T. E. (1996). *Disaster evacuation behavior: Tourists and other transients*. Boulder: University of Colorado, Institute of Behavioral Science.

- Durage, S. W., Kattan, L., Wirasinghe, S. C., & Ruwanpura, J. Y. (2014). Evacuation behaviour of households and drivers during a tornado. *Natural Hazards*, 71 (3), 1495–1517. doi:10.1007/s11069-013-0958-6.
- Echenique, M., & Melgar, L. (2018, May 11). Mapping Puerto Rico's hurricane migration with mobile phone data. *CityLab*. URL Retrieved from [https://www.citylab.com/environment/2018/05/watch-puerto-ricos-hurricane-migration-via-mobile-phone-data/559889/?utm\\_source=SFFB](https://www.citylab.com/environment/2018/05/watch-puerto-ricos-hurricane-migration-via-mobile-phone-data/559889/?utm_source=SFFB).
- Ehrenfeucht, R., & Nelson, M. (2013). Young professionals as ambivalent change agents in New Orleans after the 2005 hurricanes. *Urban Studies*, 50(4), 825-841. doi:10.1177/0042098012452323
- Elliott, J. R., & Pais, J. (2006). Race, class, and Hurricane Katrina: Social differences in human responses to disaster. *Social Science Research*, 35(2), 295-321. doi:10.1016/j.ssresearch.2006.02.003
- Ellis, M., Wright, R., & Parks, V. (2004). Work together, live apart? geographies of racial and ethnic segregation at home and at work. *Annals of the Association of American Geographers*, 94(3), 620–37. doi:10.1111/j.1467-8306.2004.00417.x.
- Ellis, R., & Levenson, E. (2017, September 8). Floridians jam highways to flee wrath of Hurricane Irma. *CNN*. Retrieved from <http://www.nbc-2.com/story/36318177/floridians-jam-highways-to-flee-wrath-of-hurricane-irma>
- Elsner, J. B., Niu, X., & Jagger, T. H. (2004). Detecting shifts in hurricane rates using a Markov chain Monte Carlo approach. *Journal of Climate*, 17(13), 2652-2666. doi:10.1175/1520-0442(2004)017<2652:dsihru>2.0.co;2
- Esnard, A. M., Sapat, A., & Mitsova, D. (2011). An index of relative displacement risk to hurricanes. *Natural Hazards*, 59(2), 833. doi:10.1007/s11069-011-9799-3
- Faggian, A., & McCann, P. (2008). Human capital, graduate migration and innovation in British regions. *Cambridge Journal of Economics*, 33(2), 317-333. doi:10.1093/cje/ben042
- Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior*, 26(2), 132-139. doi:10.1016/j.chb.2009.10.015

- Faulkner, B. (2001). Towards a framework for tourism disaster management. *Tourism Management*, 22(2), 135-147.  
doi:10.1016/S0261-5177(00)00048-0
- Finch, C., Emrich, C. T., & Cutter, S. L. (2010). Disaster disparities and differential recovery in New Orleans. *Population and Environment*, 31(4), 179-202.  
doi:10.1007/s11111-009-0099-8
- Findlay, A. M. (2011). Migrant destinations in an era of environmental change. *Global Environmental Change*, 21, S50-S58.  
doi:10.1016/j.gloenvcha.2011.09.004
- Freudiger, J., Shokri, R., & Hubaux, J. P. (2011). Evaluating the privacy risk of location-based services. In *International Conference on Financial Cryptography and Data Security* (pp. 31-46). Springer, Berlin, Heidelberg.
- Fussell, E. (2012). Space, time, and volition: Dimensions of migration theory. In Rosenblum, M. R., & Tichenor, D. J. (Eds.) *Oxford Handbook of the Politics of International Migration*, 25-52. Oxford: Oxford University Press.  
doi:10.1093/oxfordhb/9780195337228.013.0002
- Fussell, E. (2015). The long-term recovery of New Orleans' population after Hurricane Katrina. *American Behavioral Scientist*, 59(10), 1231-1245.  
doi:10.1177/0002764215591181
- Fussell, E., Curran, S. R., Dunbar, M. D., Babb, M. A., Thompson, L., & Meijer-Irons, J. (2017). Weather-related hazards and population change: A study of hurricanes and tropical storms in the United States, 1980–2012. *The Annals of the American Academy of Political and Social Science*, 669(1), 146-167.  
doi:10.1177/0002716216682942
- Fussell, E., Curtis, K. J., & DeWaard, J. (2014a). Recovery migration to the City of New Orleans after Hurricane Katrina: A migration systems approach. *Population and Environment*, 35(3), 305-322. doi:10.1007/s11111-014-0204-5
- Fussell, E., Hunter, L. M., & Gray, C. L. (2014b). Measuring the environmental dimensions of human migration: The demographer's toolkit. *Global Environmental Change*, 28, 182-191.  
doi:10.1016/j.gloenvcha.2014.07.001

- Fussell, E., Sastry, N., & VanLandingham, M. (2010). Race, socioeconomic status, and return migration to New Orleans after Hurricane Katrina. *Population and Environment*, 31(1-3), 20-42. doi:10.1007/s11111-009-0092-2
- Gall, M., Borden, K. A., Emrich, C. T., & Cutter, S. L. (2011). The unsustainable trend of natural hazard losses in the United States. *Sustainability*, 3(11), 2157-2181. doi:10.3390/su3112157
- Gerber, M. S. (2014). Predicting crime using Twitter and kernel density estimation. *Decision Support Systems*, 61, 115-125. doi:10.1016/j.dss.2014.02.003
- Girardin, F., Calabrese, F., Fiore, F. D., Ratti, C., & Blat, J. (2008). Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing*, 7(4), 36-43.
- Gladwin, H., & Peacock W. G. (1997) Warning and evacuation: A night for hard houses. In W. G. Peacock, B. Morrow, & H. Gladwin (Eds.) *Hurricane Andrew: Gender, ethnicity and the sociology of disasters* (pp. 52–74). New York: Routledge.
- Golant, S. M. (1972). The Residential Location and Spatial Behavior of the Elderly: A Canadian Example. *Vol. 143. Chicago, Illinois: University of Chicago*.
- Golledge, R. G., & Stimson, R. J. (1997). *Spatial behavior: A geographic perspective*. London: The Guilford Press.
- Gong, Y., Lin, Y., & Duan, Z. (2017). Exploring the spatiotemporal structure of dynamic urban space using metro smart card records. *Computers, Environment and Urban Systems*, 64, 169-183. doi:10.1016/j.compenvurbsys.2017.02.003
- González, M. C., Hidalgo, C. A., & Barabási, A.-L., (2008). Understanding individual human mobility patterns. *Nature*, 453, 779–782. doi:10.1038/nature06958
- Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4), 211-221. doi:10.1007/s10708-007-9111-y

- Górak, R., Luckner, M., Okulewicz, M., Porter-Sobieraj, J., & Wawrzyniak, P. (2016). Indoor localisation based on GSM signals: Multistorey building study. *Mobile Information Systems*, 2016, 2719576. doi:10.1155/2016/2719576
- Gould, K. A., & Lewis, T. L. (2007). Viewing the wreckage: Eco-disaster tourism in the wake of Katrina. *Societies Without Borders*, 2(2), 175-197. doi:10.1163/187219107X203559
- Government of Puerto Rico. (2018a). *Transformation and innovation in the wake of devastation: An economic and disaster recovery plan for Puerto Rico*. Retrieved from <http://www.p3.pr.gov/assets/pr-draft-recovery-plan-for-comment-july-9-2018.pdf>
- Government of Puerto Rico. (2018b). *Puerto Rico disaster recovery action plan*. Retrieved from [http://www.cdbg-dr.pr.gov/wp-content/uploads/2018/07/HUD-Approved-Action-Plan\\_EN.pdf](http://www.cdbg-dr.pr.gov/wp-content/uploads/2018/07/HUD-Approved-Action-Plan_EN.pdf)
- Graham, M., Poorthuis, A., & Zook, M. (2013, November 28). Digital trails of the UK floods - how well do tweets match observations? *The Guardian* Retrieved from <https://www.theguardian.com/news/datablog/2012/nov/28/data-shadows-twitter-uk-floods-mapped>
- Gray, C., & Mueller, V. (2012). Drought and population mobility in rural Ethiopia. *World Development*, 40(1), 134-145. doi:10.1016/j.worlddev.2011.05.023
- Greenwood, S., Perrin, A., & Duggan, M. (2016). *Social media update 2016*. Retrieved from [http://assets.pewresearch.org/wp-content/uploads/sites/14/2016/11/10132827/PI\\_2016.11.11\\_Social-Media-Update\\_FINAL.pdf](http://assets.pewresearch.org/wp-content/uploads/sites/14/2016/11/10132827/PI_2016.11.11_Social-Media-Update_FINAL.pdf)
- Grinberger, A. Y., Shoval, N., & McKercher, B. (2014). Typologies of tourists' time-space consumption: A new approach using GPS data and GIS tools. *Tourism Geographies*, 16(1), 105-123. doi:10.1080/14616688.2013.869249
- Groen, J. A., & Polivka, A. E. (2010). Going home after Hurricane Katrina: Determinants of return migration and changes in affected areas. *Demography*, 47(4), 821-844. doi:10.1007/BF03214587
- Gupta, A., Lamba, H., Kumaraguru, P., & Joshi, A. (2013, May). Faking Sandy: Characterizing and identifying fake images on Twitter during Hurricane Sandy. In

- Proceedings of the 22nd International Conference on World Wide Web* (pp. 729-736). ACM. Rio de Janeiro, Brazil  
doi:10.1145/2487788.2488033
- Gurley, K., Davis, R. H., Jr., Ferrera, S. P., Burton, J., Masters, F., Reinhold, T., & Abdullah, M. (2006). Post 2004 hurricane field survey--an evaluation of the relative performance of the standard building code and the Florida building code. In *Structures Congress 2006: Structural Engineering and Public Safety* (pp. 1-10). St. Louis, Missouri. doi:10.1061/40889(201)8
- Hägerstraand, T. (1970). What about people in regional science? *Papers in Regional Science Association*, 24 (January), 7–21. doi:10.1111/j.1435-5597.1970.tb01464.x.
- Hall, C. M. (2005). *Tourism: Rethinking the Social Science of Mobility*. Harloww: Prentice Hall.
- Harvey, D. (1989). *The Condition of Postmodernity. An Enquiry into the Origins of Cultural Change*. Oxford: Blackwell.
- Hasan, S., Mesa-Arango, R., & Ukkusuri, S. V. (2013a). A random-parameter hazard-based model to understand household evacuation timing behavior. *Transportation Research Part C: Emerging Technologies*, 27, 108-116. doi:10.1016/j.trc.2011.06.005
- Hasan, S., Ukkusuri, S., Gladwin, H., & Murray-Tuite, P. (2010). Behavioral model to understand household-level hurricane evacuation decision making. *Journal of Transportation Engineering*, 137(5), 341-348. doi:10.1061/(ASCE)TE.1943-5436.0000223
- Hasan, S., Zhan, X., & Ukkusuri, S. V. (2013b, August). Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. In *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing* (p. 6). ACM. Chicago, IL, USA. doi:10.1145/2505821.2505823
- Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3), 260-271. doi:10.1080/15230406.2014.890072

- Hazards Management Group, Inc. (1985). *South Carolina hurricane evacuation study. Behavioral analysis* (Report prepared for the Charleston District). U.S. Army Corps of Engineers. Tallahassee, FL: Hazards Management Group, Inc.
- Hecht, B. J., & Stephens, M. (2014). A tale of cities: Urban biases in volunteered geographic information. *ICWSM*, 14, 197-205.
- Herdağdelen, A., Adamic, L., & Mason, W. (2016, May). The social ties of immigrant communities in the United States. In *Proceedings of the 8th ACM Conference on Web Science* (pp. 78-84). ACM. Hannover, Germany. doi:10.1145/2908131.2908163
- Herfort, B., de Albuquerque, J. P., Schelhorn, S. J., & Zipf, A. (2014). Exploring the geographical relations between social media and flood phenomena to improve situational awareness. In Huerta, J., Schade, S., & Granell, C. (Eds.). *Connecting a digital Europe through location and place* (pp. 55-71). New York: Springer. doi:10.1007/978-3-319-03611-3\_4
- Hinojosa, J., & Meléndez, E. (2018). *Puerto Rican exodus: One year since Hurricane Maria*. Retrieved from Center for Puerto Rican Studies, The City University of New York website: [https://centropr.hunter.cuny.edu/sites/default/files/RB2018-05\\_SEPT2018%20%281%29.pdf](https://centropr.hunter.cuny.edu/sites/default/files/RB2018-05_SEPT2018%20%281%29.pdf)
- Hinojosa, J., Meléndez, E., & Román, N. (2018). *Puerto Rican Post-Maria relocation by states*. Retrieved from Center for Puerto Rican Studies, The City University of New York website: <https://centropr.hunter.cuny.edu/sites/default/files/PDF/Schoolenroll-v4-27-2018.pdf%0Ahttps://centropr.hunter.cuny.edu/sites/default/files/PDF/Schoolenroll-v2-3-3-2018.pdf>
- Horita, F. E., Degrossi, L. C., Assis, L. F., Zipf, A., & Albuquerque, J. P. D. (2013). The use of Volunteered Geographic Information and crowdsourcing in disaster management: A systematic literature review. In *Proceedings of the Nineteenth Americas Conference on Information Systems* (pp. 1-10). Chicago, IL, USA. Retrieved from <https://aisel.aisnet.org/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1591&context=amcis2013>

- Hsiang, S. M., & Jina, A. S. (2014). *The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones* (National Bureau of Economic Research No. w20352). doi:10.3386/w20352
- Hu, F., Li, Z., Yang, C., & Jiang, Y. (2018). A graph-based approach to detecting tourist movement patterns using social media data. *Cartography and Geographic Information Science*, 1-15.  
doi:10.1080/15230406.2018.1496036
- Hu, T., & Smith, R. B. (2018). The impact of Hurricane Maria on the vegetation of Dominica and Puerto Rico using multispectral remote sensing. *Remote Sensing*, 10(6), 827. doi:10.3390/rs10060827
- Huang, Q., & Xiao, Y. (2015). Geographic situational awareness: Mining tweets for disaster preparedness, emergency response, impact, and recovery. *ISPRS International Journal of Geo-Information*, 4(3), 1549-1568.  
<https://doi.org/10.3390/ijgi4031549>
- Huang, S. K., Lindell, M. K., & Prater, C. S. (2016a). Who leaves and who stays? A review and statistical meta-analysis of hurricane evacuation studies. *Environment and Behavior*, 48(8), 991-1029.  
doi:10.1177/0013916515578485.
- Huang, S. K., Lindell, M. K., & Prater, C. S. (2016b). Toward a multi-stage model of hurricane evacuation decision: An empirical study of Hurricanes Katrina and Rita. *Natural Hazards Review*, 18(3), 05016008 1-15. doi:10.1061/(ASCE)NH.1527-6996.0000237.
- Hughes, C., Zagheni, E., Abel, G. J., Sorichetta, A., Wisniowski, A., Weber, I., & Tatem, A. J. (2016). *Inferring migrations: Traditional methods and new approaches based on mobile phone, social media, and other Big Data* (Report prepared for the European Commission project #VT/2014/093). Retrieved from <http://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=7910&type=2&furtherPubs=yes>.
- Hunter, L. M., Luna, J. K., & Norton, R. M. (2015). Environmental dimensions of migration. *Annual Review of Sociology*, 41, 377-397.  
doi:10.1146/annurev-soc-073014-112223
- Hutton, D. (2008). *Older People in Emergencies: Considerations for Action and Policy Development*. Retrieved from The World Health Organization website: <http://www.who.int/ageing/publications/>



- Hazards & Vulnerability Research Institute (HVRI). (2017). *The Spatial Hazard Events and Losses Database for the United States*, Version 15.2. Retrieved from <http://www.sheldus.org>
- IBM. (2011). *IBM Big Data success stories*. Retrieved from <https://public.dhe.ibm.com/software/data/sw-library/big-data/ibm-big-data-success.pdf>
- Internet Live Stats (2017). *Twitter usage statistics*. Retrieved from <http://www.internetlivestats.com/one-second/#tweets-band>
- Isaacson, M., & Shoval, N. (2006). Application of tracking technologies to the study of pedestrian spatial behavior. *The Professional Geographer*, 58(2), 172-183. doi:10.1111/j.1467-9272.2006.00524.x
- Jiang, Y., Li, Z., & Ye, X. (2018). Understanding demographic and socioeconomic biases of geotagged Twitter users at the county level. *Cartography and Geographic Information Science*, 1-15. doi:10.1080/15230406.2018.1434834
- Johnson, L. A., & Hayashi, H. (2012). Synthesis efforts in disaster recovery research. *International Journal of Mass Emergencies and Disasters*, 30(2), 212-238. Retrieved from <http://www.ijmed.org/articles/602/download/>
- Johnson, T. P., & Wislar, J. S. (2012). Response rates and nonresponse errors in surveys. *Journal of the American Medical Association*, 307(17), 1805-1806. doi:10.1001/jama.2012.3532
- Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., & Newth, D. (2015). Understanding human mobility from Twitter. *PLoS ONE*, 10, e0131469. doi:10.1371/journal.pone.0131469
- Kates, R. (1962). Hazard and choice perception in flood plain management. *Department of Geography Research Paper No. 78, University of Chicago*.
- Kaufmann, V. (2017). *Re-thinking Mobility: Contemporary Sociology*. New York: Routledge. doi:10.4324/9781315244303

- Kent, J. D., & Capello, H. T., Jr. (2013). Spatial patterns and demographic indicators of effective social media content during the Horseshoe Canyon fire of 2012. *Cartography and Geographic Information Science*, 40(2), 78-89. doi:10.1080/15230406.2013.776727
- Kikas, R., Dumas, M., & Saabas, A. (2015, September). Explaining international migration in the skype network: The role of social network features. In *Proceedings of the 1st ACM Workshop on Social Media World Sensors* (pp. 17-22). ACM. Guzelyurt, Cyprus. doi:10.1145/2806655.2806658
- Kimanuka, O. (2015, May 2). Social media has changed the way we communicate. *The New Times*. Retrieved from <http://www.newtimes.co.rw/section/article/2015-05-02/188395/>
- Kishore, N., Marqués, D., Mahmud, A., Kiang, M. V., Rodriguez, I., Fuller, A., ... Buckee, C.O. (2018). Mortality in Puerto Rico after Hurricane Maria. *New England Journal of Medicine*, 379, 162-170. doi:10.1056/NEJMs1803972
- Knutson, T. R., McBride, J. L., Chan, J., Emanuel, K., Holland, G., Landsea, C., ... & Sugi, M. (2010). Tropical cyclones and climate change. *Nature Geoscience*, 3(3), 157. doi:10.1038/ngeo779
- Kontokosta, C. E., & Johnson, N. (2017). Urban phenology: Toward a real-time census of the city using Wi-Fi data. *Computers, Environment and Urban Systems*, 64, 144-153. doi:10.1016/j.compenvurbsys.2017.01.011.
- Koutsopoulos, K. C., & Schmidt, C. G. (1976). Mobility constraints of the carless. *Ergonomics*, 42, (248), 29:32
- Kryvasheyeu, Y., Chen, H., Moro, E., Van Hentenryck, P., & Cebrian, M. (2015). Performance of social network sensors during Hurricane Sandy. *PLoS ONE*, 10(2), e0117288. doi:10.1371/journal.pone.0117288
- Kryvasheyeu, Y., Chen, H., Obradovich, N., Moro, E., Van Hentenryck, P., Fowler, J., & Cebrian, M. (2016). Rapid assessment of disaster damage using social media activity. *Science Advances*, 2(3), e1500779. doi:10.1126/sciadv.1500779

- Kumar, D., & Ukkusuri, S. V. (2018, April). Utilizing geo-tagged tweets to understand evacuation dynamics during emergencies: A case study of Hurricane Sandy. *In Companion Proceedings of the Web Conference 2018* (pp. 1613-1620). Lyon, France. doi:10.1145/3184558.3191619
- Kwan, M. P. (2000). Gender differences in space-time constraints. *Area*, 32 (2), 145–56. doi:10.1111/j.1475-4762.2000.tb00125.x.
- Laczko, F. (2015). Factoring migration into the development data revolution. *Journal of International Affairs*, 68(2), 1.
- Laczko, F., & Aghazarm, C. (2009). *Migration, Environment, and Climate Change: Assessing the Evidence*. Retrieved from the International Organization for Migration website: [https://publications.iom.int/system/files/pdf/migration\\_and\\_environment.pdf](https://publications.iom.int/system/files/pdf/migration_and_environment.pdf)
- Lazo, J. K., Bostrom, A., Morss, R. E., Demuth, J. L., & Lazrus, H. (2015). Factors affecting hurricane evacuation intentions. *Risk Analysis*, 35(10), 1837-1857. doi:10.1111/risa.12407
- Leak, A., & Lansley, G. (2018). Geotemporal Twitter demographics. In Longley, P., Cheshire, J., & Singleton, A. (Eds.), *Consumer Data Research* (pp. 152–165) London: University College London Press. doi:10.2307/j.ctvqhsn6.14
- Li, J., & Rao, H. R. (2010). Twitter as a rapid response news service: An exploration in the context of the 2008 China earthquake. *The Electronic Journal of Information Systems in Developing Countries*, 42(1), 1-22. doi:10.1080/15230406.2016.1271356
- Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018a). Big data in tourism research: A literature review. *Tourism Management*, 68, 301-323. doi:10.1016/j.tourman.2018.03.009
- Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science*, 40(2), 61-77. doi:10.1080/15230406.2013.777139
- Li, Z., Wang, C., Emrich, C. T., & Guo, D. (2018b). A novel approach to leveraging social media for rapid flood mapping: A case study of the 2015 South Carolina floods.

- Cartography and Geographic Information Science*, 45(2), 97-110.  
doi:10.1080/15230406.2016.1271356
- Lindell, M. K., Kang, J. E., & Prater, C. S. (2011). The logistics of household hurricane evacuation. *Natural Hazards*, 58(3), 1093-1109.  
doi:10.1007/s11069-011-9715-x.
- Lindell, M. K., Lu, J. C., & Prater, C. S. (2005). Household decision making and evacuation in response to Hurricane Lili. *Natural Hazards Review*, 6(4), 171-179.  
[https://doi/abs/10.1061/\(ASCE\)1527-6988\(2005\)6:4\(171\)](https://doi/abs/10.1061/(ASCE)1527-6988(2005)6:4(171))
- Lund, K. A., & Benediktsson, K. (2011). Inhabiting a risky earth: The Eyjafjallajökull eruption in 2010 and its impacts. *Anthropology Today*, 27(1), 6-9.  
doi:10.1111/j.1467-8322.2011.00781.x
- Mallick, B., & Vogt, J. (2014). Population displacement after cyclone and its consequences: Empirical evidence from coastal Bangladesh. *Natural Hazards*, 73(2), 191-212. doi:10.1007/s11069-013-0803-y
- Mandel, B., Culotta, A., Boulahanis, J., Stark, D., Lewis, B., & Rodrigue, J. (2012, June). A demographic analysis of online sentiment during Hurricane Irene. In *Proceedings of the Second Workshop on Language in Social Media* (pp. 27-36). Montréal, Canada. Retrieved from <http://cs.iit.edu/~culotta/pubs/mandel12demo.pdf>
- Manovich, L. (2011). Trending: The promises and the challenges of big social data. *Debates in the Digital Humanities*, 2, 460-475.  
doi:10.5749/minnesota/9780816677948.003.0047
- Martín, Y., Li, Z., & Cutter, S. L. (2017). Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew. *PLoS ONE*, 12, e0181701. doi:10.1371/journal.pone.0181701
- Matyas, C., Srinivasan, S., Cahyanto, I., Thapa, B., Pennington-Gray, L., & Villegas, J. (2011). Risk perception and evacuation decisions of Florida tourists under hurricane threats: A stated preference analysis. *Natural Hazards*, 59(2), 871-890.  
doi:10.1007/s11069-011-9801-0

- McGranahan, G., Balk, D., & Anderson, B. (2007). The rising tide: Assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization*, 19(1), 17-37.  
doi:10.1177/0956247807076960
- McLeman, R., & Gemenne, F. (Eds.). (2018). *Routledge handbook of environmental displacement and migration*. Routledge.  
doi:10.4324/9781315638843
- McLeman, R., & Smit, B. (2006). Migration as an adaptation to climate change. *Climatic Change*, 76(1-2), 31-53. doi:10.1007/s10584-005-9000-7
- McLeman, R. A. (2014). *Climate and Human Migration: Past Experiences, Future Challenges*. Cambridge University Press.
- McLeman, R. A., & Hunter, L. M. (2010). Migration in the context of vulnerability and adaptation to climate change: Insights from analogues. *Wiley Interdisciplinary Reviews: Climate Change*, 1(3), 450-461.  
doi:10.1002/wcc.51
- Meier, P. (2015). *Digital humanitarians: How big data is changing the face of humanitarian response*. Boca Raton: CRC Press. doi:10.1201/b18023
- Melander, E., Öberg, M., & Hall, J. (2009). Are 'new wars' more atrocious? Battle severity, civilians killed and forced migration before and after the end of the Cold War. *European Journal of International Relations*, 15(3), 505-536.  
doi:10.1177/1354066109338243
- Messias, J., Benevenuto, F., Weber, I., & Zagheni, E. (2016, August). From migration corridors to clusters: The value of Google+ data for migration studies. In *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 421-428). IEEE Press. San Francisco, CA, USA.  
doi:10.1109/ASONAM.2016.7752269
- Mislove, A., Lehmann, S., Ahn, Y. Y., Onnela, J. P., & Rosenquist, J. N. (2011). Understanding the demographics of Twitter users. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. Barcelona, Spain. Retrieved from <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewFile/2816/3234>

- Mitchell, C. M., Esnard, A. M., & Sapat, A. (2012). Hurricane events, population displacement, and sheltering provision in the United States. *Natural Hazards Review*, 13(2), 150-161.  
doi:10.1061/(ASCE)NH.1527-6996.0000064
- Mitchell, J. T., Cutter, S. L., & Edmonds, A. S. (2007). Improving shadow evacuation management: Case study of the Graniteville, South Carolina chlorine spill. *Journal of Emergency Management*, 5(1), 28-34. doi:10.5055/jem.2007.0041
- Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic Perspectives*, 25(3), 173-96.  
doi:10.1257/jep.25.3.173
- Moon, H., Kim, C., & Lee, W. (2016). A UAV based 3-D positioning framework for detecting locations of buried persons in collapsed disaster area. international archives of the photogrammetry, *Remote Sensing and Spatial Information Sciences*, 41.  
doi:10.5194/isprsarchives-XLI-B8-121-2016.
- Morss, R. E., Demuth, J. L., Lazo, J. K., Dickinson, K., Lazrus, H., & Morrow, B. H. (2016). Understanding public hurricane evacuation decisions and responses to forecast and warning messages. *Weather and Forecasting*, 31(2), 395-417.  
doi:10.1175/WAF-D-15-0066.1
- Murray-Tuite, P., & Wolshon, B. (2013). Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C: Emerging Technologies*, 27, 25-45.  
<http://dx.doi.org/10.1016/j.trc.2012.11.005>
- Murthy, D., & Gross, A. J. (2017). Social media processes in disasters: Implications of emergent technology use. *Social Science Research*, 63, 356-370.  
doi:10.1016/j.ssresearch.2016.09.015
- National Hurricane Center. (2018). *Costliest U.S. tropical cyclones tables updated*. Retrieved from <https://www.nhc.noaa.gov/news/UpdatedCostliest.pdf>
- National Institute of Standards and Technology (NIST). (2017). *Hurricane Maria (2017): Preliminary Peak Wind Gust*. Retrieved from

[https://www.nist.gov/sites/default/files/documents/2018/02/20/03\\_update\\_on\\_preliminary\\_reconnaissance\\_of\\_hurricane\\_maria\\_puerto\\_rico.pdf](https://www.nist.gov/sites/default/files/documents/2018/02/20/03_update_on_preliminary_reconnaissance_of_hurricane_maria_puerto_rico.pdf)

National Oceanic and Atmospheric Administration (NOAA). (2017). *Extremely Active 2017 Atlantic Hurricane Season Finally Ends*. Retrieved from <http://www.noaa.gov/media-release/extremely-active-2017-atlantic-hurricane-season-finally-ends>

National Oceanic and Atmospheric Administration (NOAA). (2018). *State of the Climate: Hurricanes and Tropical Storms for Annual 2017*. Retrieved from <https://www.ncdc.noaa.gov/sotc/tropical-cyclones/201713>.

National Weather Service (NWS). (2017). *Major Hurricane Maria – September 20, 2017*. Retrieved from <https://www.weather.gov/sju/maria2017>

Nguyen, T. M., Kawamura, T., & Ohsuga, A. (2013). Extraction and estimation of human activity from twitter for information sharing in disaster. *Journal of Convergent Information Technology*, 8(11), 707.  
doi:10.4156/jcit.vol8.issue11.79

Oliver-Smith, A. (2009). Development-forced displacement and resettlement: A global human rights crisis. In Oliver-Smith, A. (Ed.). *Development and Dispossession: The crisis of forced displacement and resettlement*, (pp. 3–23). Santa Fe: School for Advanced Research Press.

Page, S. J., & Hall, C. M. (2014). *The Geography of Tourism and Recreation: Environment, Place and Space*. New York: Routledge.

Pais, J. F., & Elliott, J. R. (2008). Places as recovery machines: Vulnerability and neighborhood change after major hurricanes. *Social Forces*, 86(4), 1415-1453.  
doi:10.1353/sof.0.0047

Palen, L., Vieweg, S., Liu, S. B., & Hughes, A. L. (2009). Crisis in a networked world: Features of computer-mediated communication in the April 16, 2007, Virginia Tech event., *Social Science Computer Review*, 27(4), 467-480.  
doi:10.1177/0894439309332302

- Pasch, R., Penny, A., & Berg, R. (2018). *Hurricane Maria (AL 152017)* Retrieved from National Oceanic and Atmospheric Administration, National Hurricane Center website: [https://www.nhc.noaa.gov/data/tcr/AL152017\\_Maria.pdf](https://www.nhc.noaa.gov/data/tcr/AL152017_Maria.pdf)
- Perttunen, M., Kostakos, V., Riekki, J., & Ojala, T. (2014). Spatio-temporal patterns link your digital identities. *Computers, Environment and Urban Systems*, 47, 58–67. doi:10.1016/j.compenvurbsys.2013.12.004
- Pettiway, L. E. (1995). Copping Crack: The travel behavior of crack users. *Justice Quarterly*, 12 (3), 499–524. doi:10.1080/07418829500096111
- Pezzullo, P. C. (2009). Tourists and/as disasters: Rebuilding, remembering, and responsibility in New Orleans. *Tourist Studies*, 9(1), 23-41. doi:10.1177/1468797609360591
- Phillips, B. D., & Morrow, B. H. (2007). Social science research needs: Focus on vulnerable populations, forecasting, & warnings. *Natural Hazards Review*, 8(3), 61-68. doi:10.1061/(asce)1527-6988(2007)8:3(61)
- Phithakkitnukoon, S., Smoreda, Z., & Olivier, P. (2012). Socio-geography of human mobility: A study using longitudinal mobile phone data. *PloS one*, 7(6), e39253. doi:10.1371/journal.pone.0039253.
- Piguet, E. (2010). Linking climate change, environmental degradation, and migration: A methodological overview. *Wiley Interdisciplinary Reviews: Climate Change*, 1(4), 517-524. doi:10.1002/wcc.54
- Piguet, E., Pécoud, A., & De Guchteneire, P. (2011). Migration and climate change: An overview. *Refugee Survey Quarterly*, 30(3), 1-23. doi:10.1093/rsq/hdr006
- Pred, A. (1981). Social reproduction and the time-geography of everyday life. *Geografiska Annaler. Series B, Human Geography* 63(1), 5-22. doi:10.2307/490994
- Quiñones-Pérez, A. T., & Seda-Irizarry, I. J. (2016). Wealth extraction, governmental servitude, and social disintegration in colonial Puerto Rico. *New Politics*, 15(4), 91.



- Rango, M., and Vespe, M. (2017). *Big Data and Alternative Data Sources on Migration: From Case-Studies to Policy Support* (Summary report European Commission – Joint Research Centre) Retrieved from <https://gmdac.iom.int/big-data-and-alternative-data-sources-on-migration-from-case-studies-to-policy-support>
- Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, 57, 202-212.  
doi:10.1016/j.tourman.2016.06.006
- Ren, F., Tong, D., & Kwan, M. P. (2014). Space-time measures of demand for service: Bridging location modelling and accessibility studies through a time-geographic framework. *Geografiska Annaler, Series B: Human Geography*, 96(4), 329–44.  
doi:10.1111/geob.12055
- Ritchie, B. (2008). Tourism disaster planning and management: From response and recovery to reduction and readiness. *Current Issues in Tourism*, 11(4), 315-348.  
doi:10.1080/13683500802140372
- Roick, O., & Heuser, S. (2013). Location based social networks—definition, current state of the art and research agenda. *Transactions in GIS*, 17(5), 763-784.  
doi:10.1111/tgis.12032
- Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213), 1063-1064. doi:10.1126/science.346.6213.1063
- Sadri, A. M., Hasan, S., & Ukkusuri, S. V. (2017). Joint inference of user community and interest patterns in social interaction networks. *Social Network Analysis and Mining*, 9(1), 11. doi: 10.1007/s13278-019-0551-4
- Sadri, A. M., Hasan, S., Ukkusuri, S. V., & Cebrian, M. (2017c). Understanding information spreading in social media during Hurricane Sandy: User activity and network properties. *arXiv preprint arXiv:1706.03019*. Retrieved from <http://arxiv.org/abs/1706.03019>
- Sadri, A. M., Ukkusuri, S. V., & Gladwin, H. (2017a). Modeling joint evacuation decisions in social networks: The case of Hurricane Sandy. *Journal of Choice Modelling*, 25, 50-60.  
doi:10.1016/j.jocm.2017.02.002

- Sadri, A. M., Ukkusuri, S. V., & Gladwin, H. (2017b). The role of social networks and information sources on hurricane evacuation decision making. *Natural Hazards Review*, 18(3), 04017005.  
doi:10.1061/(ASCE)NH.1527-6996.0000244
- Sadri, A. M., Ukkusuri, S. V., & Murray-Tuite, P. (2013). A random parameter ordered probit model to understand the mobilization time during hurricane evacuation. *Transportation Research Part C: Emerging Technologies*, 32, 21-30.  
doi:10.1016/j.trc.2013.03.009
- Sadri, A. M., Ukkusuri, S. V., Murray-Tuite, P., & Gladwin, H. (2014a). How to evacuate: Model for understanding the routing strategies during hurricane evacuation. *Journal of Transportation Engineering*, 140(1), 61-69.  
doi:10.1061/(ASCE)TE.1943-5436.0000613
- Sadri, A. M., Ukkusuri, S. V., Murray-Tuite, P., & Gladwin, H. (2014b). Analysis of hurricane evacuee mode choice behavior. *Transportation Research Part C: Emerging Technologies*, 48, 37-46.  
doi:10.1016/j.trc.2014.08.008
- Sadri, A. M., Ukkusuri, S. V., Murray-Tuite, P., & Gladwin, H. (2015). Hurricane evacuation route choice of major bridges in Miami Beach, Florida. *Transportation Research Record: Journal of the Transportation Research Board*, 2532(1), 164-173. doi:10.3141/2532-18
- Sakaki, T., Okazaki, M., & Matsuo, Y. (2010, April). Earthquake shakes Twitter users: Real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web* (pp. 851-860). ACM. Raleigh, NC, USA. doi:10.1145/1772690.1772777
- Sakaki, T., Toriumi, F., & Matsuo, Y. (2011, December). Tweet trend analysis in an emergency situation. In *Proceedings of the Special Workshop on Internet and Disasters* (p. 3). ACM. Tokyo, Japan.  
doi:10.1145/2079360.2079363
- San, Y., Clarence, S., Iii, W., Thorkildsen, Z., & Giovachino, M. (2013). *Social Media in the Emergency Management Field. 2012 Survey Results* (IPP-2013-U-004984/Final). Retrieved from [http://cdrmaguire.com/emkey/Resources/Social/SocialMedia\\_EmergencyManagement.pdf](http://cdrmaguire.com/emkey/Resources/Social/SocialMedia_EmergencyManagement.pdf)

- San Carlos Arce, R., Onuki, M., Esteban, M., & Shibayama, T. (2017). Risk awareness and intended tsunami evacuation behaviour of international tourists in Kamakura City, Japan. *International Journal of Disaster Risk Reduction*, 23. 178–92. doi:10.1016/j.ijdr.2017.04.005.
- Santiago, L., Hosseini, B., & Ellis, R. (2017, December 9). Puerto Rico: Death toll from hurricane climbs to 64 with 2 'indirect deaths'. *CNN*. Retrieved from <https://www.cnn.com/2017/12/09/us/puerto-rico-hurricane-deaths-and-assistance/index.html>.
- Sasaki, K., & Nishii, K. (2010). Measurement of intention to travel considering the effect of telecommunications on trips. *Transportation Research Part C: Emerging Technologies*, 18(1), 36–44. doi:10.1016/j.trc.2009.04.012
- Sather-Wagstaff, J. (2011). *Heritage that Hurts: Tourists in the Memoryscapes of September 11*. Walnut Creek, CA: Left Coast Press. doi:10.4324/9781315427539
- Shareck, M., Kestens, Y., & Gauvin, L. (2013). Examining the spatial congruence between data obtained with a novel activity location questionnaire, continuous GPS tracking, & prompted recall surveys. *International Journal of Health Geographics*, 12, 40. doi:10.1186/1476-072X-12-40
- Sheller, M., & Urry, J. (2006). The new mobilities paradigm. *Environment and Planning A*, 38(2), 207–226. doi:10.1068/a37268
- Shelton, T., Poorthuis, A., Graham, M., & Zook, M. (2014). Mapping the data shadows of Hurricane Sandy: Uncovering the sociospatial dimensions of 'big data'. *Geoforum*, 52, 167–179. doi:10.1016/j.geoforum.2014.01.006
- Shepherd, K. (2017, September 7). Georgia governor orders mandatory evacuation for areas of the state ahead of Irma. *WGXA*. Retrieved from <https://wgxa.tv/news/local/georgia-governor-orders-mandatory-evacuation-for-areas-of-the-state-ahead-of-irma>
- Shih, T. H., & Fan, X. (2009). Comparing response rates in e-mail and paper surveys: A meta-analysis. *Educational Research Review*, 4(1), 26–40. <http://dx.doi.org/10.1016/j.edurev.2008.01.003>
- Shoval, N., & Ahas, R. (2016). The use of tracking technologies in tourism research: The first decade. *Tourism Geographies*, 18(5), 587–606.

doi:10.1080/14616688.2016.1214977

- Shoval, N., McKercher, B., Ng, E., & Birenboim, A. (2011). Hotel location and tourist activity in cities. *Annals of Tourism Research*, 38(4), 1594-1612.  
doi:10.1016/j.annals.2011.02.007
- Siebeneck, L. K., & Cova, T. J. (2012). Spatial and temporal variation in evacuee risk perception throughout the evacuation and return-entry process. *Risk Analysis*, 32(9), 1468–80. doi:10.1111/j.1539-6924.2011.01781.x.
- Sinclair, M., O'Toole, J., Malawaraarachchi, M., & Leder, K. (2012). Comparison of response rates and cost-effectiveness for a community-based survey: Postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Medical Research Methodology*, 12(1), 132. doi:10.1186/1471-2288-12-132
- Slovic, P. (1987). Perception of risk. *Science*, 236(4799), 280–85.  
doi:10.1126/science.3563507.
- Smith, A. B., & Katz, R. W. (2013). US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. *Natural Hazards*, 67(2), 387-410.  
doi:10.1007/s11069-013-0566-5
- Smith, S. K., & McCarty, C. (2009). Fleeing the storm(s): An examination of evacuation behavior during Florida's 2004 hurricane season. *Demography*, 46, 127–145.  
doi:10.1353/dem.0.0048
- Smyth, J. D., Dillman, D. A., Christian, L. M., & O'Neill, A. C. (2010). Using the Internet to survey small towns and communities: Limitations and possibilities in the early 21st century. *American Behavioral Scientist*, 53(9), 1423-1448.  
doi:10.1177/0002764210361695
- Snijders, C., Matzat, U., & Reips, U. D. (2012). "Big data": Big gaps of knowledge in the field of internet science. *International Journal of Internet Science*, 7(1), 1-5.
- Sobolevsky, S., Sitko, I., Des Combes, R. T., Hawelka, B., Arias, J. M., & Ratti, C. (2014). Money on the move: Big data of bank card transactions as the new proxy for human mobility patterns and regional delineation. the case of residents and foreign visitors in Spain. In *Proceedings of the 2014 IEEE International Congress on Big Data* (pp. 136-143). IEEE. Washington, DC, USA.  
doi:10.1109/BigData.Congress.2014.28.

- Sorensen, J. H., & Mileti, D. S. (1988). Warning and evacuation: Answering some basic questions. *Industrial Crisis Quarterly*, 2, 195–209.  
doi:10.1177/108602668800200302
- South Carolina Department of Transportation (SCDOT). (2017). *Average annual daily traffic*. Retrieved from <https://www.scdot.org/default.aspx>
- Spence, P. R., Lachlan, K. A., & Raine, A. M. (2016). Social media and crisis research: Data collection and directions. *Computers in Human Behavior*, 54, 667-672.  
doi:10.1016/j.chb.2015.08.045
- Spencer, N., & Urquhart, M. A. (2018). Hurricane strikes and migration: Evidence from storms in Central America and the Caribbean. *Weather, Climate, and Society*, 10(3), 569-577. doi:10.1175/WCAS-D-17-0057.1
- Spyratos, S., Vespe, M., Natale, F., Ingmar, W., Zagheni, E., & Rango, M. (2018). *Migration data using social media: A European perspective* (EUR 29273 EN).  
doi:10.2760/964282.
- Squire, V. (Ed.). (2010). *The Contested Politics of Mobility: Borderzones and Irregularity*. New York: Routledge.
- Stanton, M. C., Yamauchi, M., Mkwanda, S. Z., Ndhlovu, P., Matipula, D. E., Mackenzie, C., & Kelly-Hope, L. A. (2017). Measuring the physical and economic impact of filarial lymphoedema in Chikwawa district, Malawi: A case-control study. *Infectious Diseases of Poverty*, 6(1), 28. doi:10.1186/s40249-017-0241-2
- State, B., Rodriguez, M., Helbing, D., & Zagheni, E. (2014) Migration of Professionals to the U.S. In: Aiello L.M., & McFarland D. (Eds.) *Social Informatics*. New York: Springer. doi:10.1007/978-3-319-13734-6\_37
- State, B., Weber, I., & Zagheni, E. (2013, February). Studying inter-national mobility through IP geolocation. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*. (pp. 265-274). ACM. Rome, Italy.  
doi:10.1145/2433396.2433432
- Steiger, E., De Albuquerque, J. P., & Zipf, A. (2015). An advanced systematic literature review on spatiotemporal analyses of Twitter data. *Transactions in GIS*, 19(6), 809-834. doi:10.1111/tgis.12132

- Stewart, S. R. (2017). *Hurricane Matthew (AL142016)*. Retrieved from National Oceanic and Atmospheric Administration, National Hurricane Center website: [http://www.nhc.noaa.gov/data/tcr/AL142016\\_Matthew.pdf](http://www.nhc.noaa.gov/data/tcr/AL142016_Matthew.pdf)
- Stock, K. (2018). Mining location from social media: A systematic review. *Computers, Environment and Urban Systems*, 71, 209-240  
doi:10.1016/j.compenvurbsys.2018.05.007
- Stone, L. (2017). *How low will Puerto Rico's population go?* Retrieved from <https://medium.com/migration-issues/how-low-will-puerto-ricos-population-go-c8d108ac8b3b>
- Sui, D., & Goodchild, M. (2011). The convergence of GIS and social media: Challenges for GIScience. *International Journal of Geographical Information Science*, 25(11), 1737-1748.  
doi:10.1080/13658816.2011.604636
- Sutter, J. D., & Hernandez, S. (2018, February 21). 'Exodus' from Puerto Rico: A visual guide. *CNN*. Retrieved from <https://www.cnn.com/2018/02/21/us/puerto-rico-migration-data-invs/index.html>
- Tamima, U., & Chouinard, L. (2016). Development of evacuation models for moderate seismic zones: A case study of Montreal. *International Journal of Disaster Risk Reduction*, 16, 167–79.  
doi:10.1016/j.ijdr.2016.02.003.
- Tao, S., Rohde, D., & Corcoran, J. (2014). Examining the spatial–temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap. *Journal of Transport Geography*, 41, 21-36.  
doi:10.1016/j.jtrangeo.2014.08.006
- Taylor, L. (2016). No place to hide? The ethics and analytics of tracking mobility using mobile phone data. *Environment and Planning D: Society and Space*, 34(2), 319-336. doi:10.1177/0263775815608851
- Teralytics (2018). Mapping Puerto Rico's hurricane migration with mobile phone data. *Citylab*. Retrieved from <https://www.citylab.com/environment/2018/05/watch-puerto-ricos-hurricane-migration-via-mobile-phone-data/559889/>.

- Thinus-Blanc, C., & Gaunet, F. (1997). Representation of space in blind persons: Vision as a spatial sense? *Psychological Bulletin*, 121(1), 20–42. doi:10.1037/0033-2909.121.1.20
- Toole, M. J. (1995). Mass population displacement. A global public health challenge. *Infectious Disease Clinics of North America*, 9(2), 353-366.
- Twitter. (2016). *Twitter Usage/Company Facts*. Retrieved from <https://about.twitter.com/es/company>
- Ukkusuri, S. V., Zhan, X., Sadri, A. M., & Ye, Q. (2014). Use of social media data to explore crisis informatics: Study of 2013 Oklahoma Tornado. *Transportation Research Record*, 2459(1), 110-118. doi:10.3141/2459-13
- United Nations (2008), *Principles and Recommendations for Population and Housing Censuses* (Statistical Papers (Ser. M)). doi:10.18356/be1ae14b-en.
- United Nations (2014). *A world that counts: Mobilising the data revolution for sustainable development* (UN IEAG, 2014). Retrieved from <http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>
- United States Census Bureau (2018a). *December 19, 2018. Release number CB18-193*. Retrieved from <https://www.census.gov/newsroom/press-releases/2018/estimates-national-state.html>
- United States Census Bureau (2018b). *2017 American Community Survey*. Retrieved from <https://www.census.gov/programs-surveys/acs>
- Van Acker, V., Van Wee, B., & Witlox, F. (2010). When transport geography meets social psychology: Toward a conceptual model of travel behaviour. *Transport Reviews*, 30(2), 219-240. doi:10.1080/01441640902943453
- Van Willigen, M., Edwards, B., Lormand, S., & Wilson, K. (2005). Comparative assessment of impacts and recovery from Hurricane Floyd among student and community households. *Natural Hazards Review*, 6(4), 180-190. doi:10.1061/(ASCE)1527-6988(2005)6:4(180)

- Vazquez-Prokopec, G. M., Stoddard, S. T., Paz-Soldan, V., Morrison, A. C., Elder, J. P., Kochel, T. J., Scott, T.W., & Kitron, U. (2009). Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. *International Journal of Health Geographics*, 8(1), 68. doi:10.1186/1476-072x-8-68
- Versichele, M., De Groote, L., Bouuaert, M. C., Neutens, T., Moerman, I., & Van de Weghe, N. (2014). Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. *Tourism Management*, 44, 67-81  
doi:10.1016/j.tourman.2014.02.009
- Villegas, J., Matyas, C., Srinivasan, S., Cahyanto, I., Thapa, B., & Pennington-Gray, L. (2013). Cognitive and affective responses of Florida tourists after exposure to hurricane warning messages. *Natural Hazards*, 66(1), 97-116.  
doi:10.1007/s11069-012-0119-3
- Vogel, N., Theisen, C., Leidig, J. P., Scripps, J., Graham, D. H., & Wolffe, G. (2015). mining mobile datasets to enable the fine-grained stochastic simulation of ebola diffusion. *Procedia Computer Science*, 51, 765-774.  
doi:10.1016/j.procs.2015.05.197.
- Walters, W. (2006). Border/control. *European Journal of Social Theory*, 9(2), 187-203.  
doi:10.1177/1368431006063332
- Wang, H., Zhang, X., Wu, L., Hou, C., Gong, H., Zhang, Q., & Ouyang, M. (2015). Beijing passenger car travel survey: Implications for alternative fuel vehicle deployment. *Mitigation and Adaptation Strategies for Global Change*, 20(5), 817-835. doi:10.1007/s11027-014-9609-9
- Wang, Q., & Taylor, J. E. (2014). Quantifying, comparing human mobility perturbation during Hurricane Sandy, Typhoon Wipha, Typhoon Haiyan. *Procedia Economics and Finance*, 18, 33-38.  
doi:10.1016/S2212-5671(14)00910-1
- Wang, Q., & Taylor, J. E. (2016). Patterns and limitations of urban human mobility resilience under the influence of multiple types of natural disaster. *PLoS ONE*, 11(1), e0147299. doi:10.1371/journal.pone.0147299



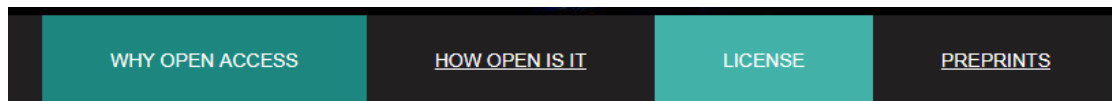
- Wesolowski, A., Buckee, C. O., Bengtsson, L., Wetter, E., Lu, X., & Tatem, A. J. (2014). Commentary: Containing the ebola outbreak-the potential and challenge of mobile network data. *PLOS Currents*, 6. doi:10.1371/currents.outbreaks.0177e7fcf52217b8b634376e2f3efc5e
- Wesolowski, A., Buckee, C. O., Engø-Monsen, K., & Metcalf, C. J. E. (2016). Connecting mobility to infectious diseases: The promise and limits of mobile phone data. *The Journal of Infectious Diseases*, 214(suppl\_4), S414-S420. doi:10.1093/infdis/jiw273.
- Wesolowski, A., Eagle, N., Tatem, A. J., Smith, D. L., Noor, A. M., Snow, R. W., & Buckee, C. O. (2012). Quantifying the impact of human mobility on malaria. *Science*, 338(6104), 267-270. doi:10.1126/science.1223467
- White, G. F. (1945). Human adjustments to floods: A geographical approach to the flood problem in the United States. *Research Paper No. 29. Chicago: Department of Geography, University of Chicago*. Retrieved from [https://biotech.law.lsu.edu/climate/docs/Human\\_Adj\\_Floods\\_White.pdf](https://biotech.law.lsu.edu/climate/docs/Human_Adj_Floods_White.pdf)
- Whitehead, J. C., Edwards, B., Van Willigen, M., Maiolo, J. R., Wilson, K., & Smith, K. T. (2000). Heading for higher ground: Factors affecting real and hypothetical hurricane evacuation behavior. *Global Environmental Change Part B: Environmental Hazards*, 2(4), 133-142. doi:10.3763/ehaz.2000.0219
- Widener, M. J., & Li, W. (2014). Using geolocated Twitter data to monitor the prevalence of healthy and unhealthy food references across the US. *Applied Geography*, 54, 189-197. doi:10.1016/j.apgeog.2014.07.017
- Widhalm, P., Yang, Y., Ulm, M., Athavale, S., & González, M. C. (2015). Discovering urban activity patterns in cell phone data. *Transportation*, 42(4), 597-623. doi:10.1007/s11116-015-9598-x
- Wilks, A. G. (2017, September 8). Governor orders evacuations of SC barrier islands as Irma approaches. *The State*. Retrieved from <https://www.thestate.com/news/politics-government/article172169352.html>

- Willekens, F., Massey, D., Raymer, J., & Beauchemin, C. (2016). International migration under the microscope. *Science*, 352(6288), 897-899. doi:10.1126/science.aaf6545
- Williams, N. E., Thomas, T. A., Dunbar, M., Eagle, N., & Dobra, A. (2015). Measures of human mobility using mobile phone records enhanced with GIS data. *PloS ONE*, 10(7), e0133630. doi:10.1371/journal.pone.0133630
- Willoughby, H. E. (2012). Distributions and trends of death and destruction from hurricanes in the United States, 1900–2008. *Natural Hazards Review*, 13(1), 57-64. doi:10.1061/(ASCE)NH.1527-6996.0000046
- Wilson, S. N., & Tiefenbacher, J. P. (2012). The barriers impeding precautionary behaviours by undocumented immigrants in emergencies: The Hurricane Ike experience in Houston, Texas, USA. *Environmental Hazards*, 11(3), 194-212. doi:10.1080/17477891.2011.649711
- Wolshon, B., & Pande, A. (2016). *Traffic Engineering Handbook*. John Wiley & Sons.
- Wu, H. C., Lindell, M. K., & Prater, C. S. (2012). Logistics of hurricane evacuation in Hurricanes Katrina and Rita. *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(4), 445-461. doi:10.1016/j.trf.2012.03.005
- Wukich, C. (2015). Social media use in emergency management. *Journal of Emergency Management*, 13(4), 281-294. doi:10.5055/jem.2015.0242
- Yan, Y., Eckle, M., Kuo, C. L., Herfort, B., Fan, H., & Zipf, A. (2017). Monitoring and assessing post-disaster tourism recovery using geotagged social media data. *ISPRS International Journal of Geo-Information*, 6(5), 144. doi:10.3390/ijgi6050144
- Zagheni, E., Garimella, V. R. K., & Weber, I. (2014, April). Inferring international and internal migration patterns from twitter data. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 439-444). ACM. Seoul, Republic of Korea doi:10.1145/2567948.2576930
- Zagheni, E., Polimis, K., Alexander, M., Weber, I., & Billari, F. C. (2018). Combining social media data and traditional surveys to nowcast migration stocks. In *Proceedings of the Annual Meeting of the Population Association of America*. Austin, TX, USA.

- Zagheni, E., & Weber, I. (2012, June). You are where you e-mail: Using e-mail data to estimate international migration rates. In *Proceedings of the 4th Annual ACM Web Science Conference* (pp. 348-351). ACM. Evanston, IL, USA. doi:10.1145/2380718.2380764
- Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook's advertising platform to monitor stocks of migrants. *Population and Development Review*, 43(4), 721-734. doi:10.1111/padr.12102
- Zeigler, D. J., Brunn, S. D., & Johnson, J. H., Jr. (1981). Evacuation from a nuclear technological disaster. *Geographical Review*, 71, 1-16. doi:10.2307/214548
- Zhai, S., Xu, X., Yang, L., Zhou, M., Zhang, L., & Qiu, B. (2015). Mapping the popularity of urban restaurants using social media data. *Applied Geography*, 63, 113-120. doi:10.1016/j.apgeog.2015.06.006
- Zhang, Y., Prater, C. S., & Lindell, M. K. (2004). Risk area accuracy and evacuation from Hurricane Bret. *Natural Hazards Review*, 5(3), 115-120. doi:10.1061/(asce)1527-6988(2004)5:3(115)
- Zhong, C., Manley, E., Arisona, S. M., Batty, M., & Schmitt, G. (2015). Measuring variability of mobility patterns from multiday smart-card data. *Journal of Computational Science*, 9, 125-130. doi:10.1016/j.jocs.2015.04.021

## APPENDIX A

### PLOS ONE LICENSE



## License

PLOS applies the [Creative Commons Attribution](#) (CC BY) license to works we publish. Under this license, authors retain ownership of the copyright for their content, but they allow anyone to download, reuse, reprint, modify, distribute and/or copy the content as long as the original authors and source are cited.

Appropriate attribution can be provided by simply citing the original article (e.g., Huntingtin Interacting Proteins Are Genetic Modifiers of Neurodegeneration. Kaltenbach LS et al. *PLOS Genetics*. 2007. 3(5) doi:10.1371/journal.pgen.0030082).

If you have a question about the license, please [email us](#).

Figure A.1. Permission to reprint: Plos ONE license. Source:  
<https://www.plos.org/license>